A DEMOGRAPHIC AND DIETARY HISTORY OF ANCIENT DOGS IN THE AMERICAS USING ANCIENT DNA

BY

KELSEY ELISSA WITT DILLON

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Ecology, Evolution, and Conservation Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Associate Professor Ripan S. Malhi, Chair and Director of Research
Assistant Professor Anna V. Kukekova
Associate Professor Alfred L. Roca
Professor Stanley H. Ambrose
Associate Professor Brian M. Kemp, University of Oklahoma

ABSTRACT

Dogs were domesticated more than 15,000 years ago, and since then they have become an integral part of human lives. They have served as hunters, guards, and pets, and have migrated with humans to multiple continents, including the Americas and Australia. The close relationship between humans and dogs makes dogs a valuable proxy when studying human history. In this study, we use ancient dog remains from the Americas to gain an understanding of their demographic and dietary history, as well as that of humans. Mitochondrial DNA sequences of the hypervariable region of ancient dogs were compared to modern and ancient American dogs to model dog demography and compare populations to identify shared haplotypes. This study identified multiple founding haplotypes, and suggested that dogs arrived to the Americas after the initial human migration. The majority of published ancient American dog DNA sequences is of the hypervariable region, so this comparison gives us the opportunity to look at the largest number of dogs across the Americas. We also sequenced complete mitochondrial genomes (mitogenomes), to determine if mitogenome data could be used to confirm the hypotheses made about ancient American dog demography using the hypervariable region. Mitogenome sequences show a higher-resolution perspective on dog diversity, and the longer sequences revealed different aspects of dog demography. We were able to support the hypotheses that suggest that dogs migrated to the Americas with humans, and that dog populations vary in genetic diversity, but were not able to support the hypotheses that ancient and modern dogs show continuity, and that dogs arrived to the Americas later in time. We also found that ancient dog demography mirrors ancient Native American demography in specific regions of North America, such as the Pacific Coast and Southeast. Finally, we assessed the diet in dogs from the American Bottom using both stable isotopes and shotgun sequencing of dog coprolites, and used the findings about dog diet to infer human diet during the Late Woodland and Mississippian periods. We found that dogs (and humans) ate no maize during the Late Woodland Period, but were consuming large amounts of maize as early as 1010 AD, and maize was likely present in the American

Bottom by 900 AD. Additionally, Mississippian dogs and humans supplemented their diet of maize with other foods including squash and fish. The analysis of the history of dogs has yielded a wealth of information about how dogs and humans interacted in the Americas.

ACKNOWLEDGMENTS

I would like to thank my advisor Ripan Malhi for his support throughout this project, and for his advice regarding my research and the preparation of this dissertation. I would also like to thank my other committee members, Al Roca, Anna Kukekova, Brian Kemp, and Stan Ambrose, for their guidance throughout this project.

I would also like to thank the other members of the Malhi Molecular Anthropology Lab, both past and present. John Lindo and Liz Mallott were instrumental in training me in ancient DNA and bioinformatics techniques, and I owe much to their expertise. I think Cris Hughes for always providing insightful feedback on my writing and presentations. I'd like to also thank my fellow graduate students (Amanda Owings, Mary Rogers, Alyssa Bader, and Karthik Yarlagadda) as well as Malhi lab postdocs (Charla Marshall, Shizhu Gao, and Hongjie Li) for supporting my research, advising me on my work, and for providing a great social support network outside of work.

I owe much to the Illinois State Archaeological Survey (ISAS), which has been an enormously helpful resource throughout my graduate career. I thank Tom Emerson for being supportive of and enthusiastic about ancient dog research, and Kris Hedman for being my liason to ISAS and helping me with whatever was needed. I thank Eve Hargrave, Steven Kuehn, and Mary Simon for sharing their archaeological expertise, as well as everyone at ISAS for allowing me to work with their samples, providing me with project funding, and helping me to put my research into an archaeological context. I'd also like to thank other individuals who have shared their samples with me, including Brian Kemp, John Johnson, Kelsey Noack Myers, Liz Watts Malouchos, Rika Kaestle, Marilyn Masson, Eske Willerslev, and Greger Larson.

Some of my research was conducted at the Copenhagen Centre for Geogenetics, and I thank Eske Willerslev and Tom Gilbert for hosting me. I also would like to thank Tom's postdocs and graduate students, who took the time to train me in new techniques and help me acclimatize to the lab, including

Nathan Wales, Inge Lundstroem, and Marcela Sandoval Velasco.

I would like to acknowledge the Roy J. Carver Biotechnology Center for their sequencing expertise – nearly all of the sequences discussed here were sequenced at the Biotechnology Center. For troubleshooting next-generation sequencing, I would like to thank Chris Fields and Alvaro Hernandez. I had assistance with processing the sequencing results, and would like to acknowledge Julie Allen and Chris Widga and hpcbio for helping me construct a bioinformatics pipeline for processing my coprolite data.

I have had multiple funding sources, and I would like to acknowledge them for enabling me to perform my research. I received an NSF Doctoral Dissertation Research Improvement Grant (NSF BCS-1540336) and a Wenner Gren Dissertation Fieldwork Grant (Gr. 9254). Smaller research grants were funded by the Illinois State Archaeological Survey Ancient Technologies and Archaeological Materials program, and by the Program of Ecology, Evolution, and Conservation Biology at the University of Illinois at Urbana-Champaign.

Finally, I would like to thank my friends and family. My graduate school friends, including Amanda Owings, Selina Ruzi, Nicholas Sly, Cassie Wesseln, Lorena Rios, Miles Bensky, Halie Rando, Jessica Hekman, Tolu Perrin-Stowe, Alida deFlamingh, and Hannah Wahl were all great supporters of me both academically and personally, and I will always be grateful for our game nights, lunch dates, and other adventures in Illinois. I'd also like to thank my parents, Joel Witt and Holly Hunter, for their constant support of my work (and my move from Texas to Illinois), as well as my twin sister Lindsey, who has always been there when I needed it. My husband Brad I'd like to thank especially, for being such a great cheerleader and supporter of my work. I am grateful for all of the dinner dates, gaming adventures, and even troubleshooting of my code – I truly could not have done it without you, sweetheart. And last, but certainly not least, I'd like to thank Sophie for being such an excellent grad school dog, a friend and family member for 11 years, and a reminder of why this research is so worthwhile.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION

Dogs and humans have shared a close relationship for thousands of years. Dogs were one of the first species to be domesticated and have traveled widely with humans as they peopled the world, even to Australia and the Americas (Leonard et al., 2002; Savolainen et al., 2004; Greig et al., 2015; Witt et al., 2015). Because of this close relationship, dogs and humans have a shared history, and have been shown to adapt to changes in environment and lifestyle in similar ways (Axelsson et al., 2013; Li et al., 2014). Dogs can be used as a proxy to study human history, and this is particularly useful in the Americas, where dogs were abundant and utilized by many peoples for thousands of years (Schwartz, 1997). Additionally, given the ethical concerns that sometimes accompany the analysis of human remains, the study of ancient dogs can be a way to learn about human history in the Americas while still respecting the wishes of modern descendants of ancient humans.

Objectives

The primary objective of this research is to use ancient DNA techniques to clarify the demographic history of dogs in the Americas, from the timing of their entry to the Americas to the present. Mitochondrial DNA (both in part and in whole) was sequenced from multiple populations and time periods and compared to assess both levels of diversity and shared lineages, to infer how dogs were used in different populations and whether dog populations were continuous or experienced replacement through time. Understanding of how dog populations changed over time can help us infer how human demography has changed over time as well. The demographic history of dogs can also be used to reveal aspects of human culture. For example, shared lineages between dog populations could indicate migration or trade interactions. Low levels of genetic diversity could be indications that dogs were deliberately being bred. Also, the burial context in which dogs were found can also inform human cultural practices from the same time period. The use of complete mitochondrial genome (mitogenome) sequences is fairly novel in the Americas, and these mitogenomes can be used to test hypotheses about

dog demography in the Americas that were identified using shorter mitochondrial DNA sequences.

A secondary objective focuses on the use of dogs as a dietary proxy for humans to assess the arrival of maize to southern Illinois, which became the center of a large agricultural empire known as the Mississippians around 1000 years before present (ybp). The timing of maize arrival was estimated using stable isotope analysis of dog bones and teeth (specifically focusing on ^{13}C, which distinguishes between different types of plants and ^{15}N, which can distinguish between carnivore and herbivore diets), as well as shotgun sequencing and taxonomic analysis of dog coprolites. Human remains from this period of transition to maize agriculture in Illinois are unavailable for study, so dogs can be useful in pinpointing when maize arrived to the region.

Dog Domestication

Dogs were the first animals to be domesticated, and hold a unique position in human lives. They are known to have been domesticated from the gray wolf (Clutton-Brock, 1995), but the timing and origin of dog domestication is still unresolved. Using various molecular clocks, dog domestication likely occurred anywhere from 21,000 years before present (ybp) to 15,000 ybp (Pang et al., 2009; Sacks et al., 2013; Skoglund et al., 2015). However, ancient canids have been found that date in excess of 30,000 ybp, and have features similar to modern dogs, suggesting that domestication occurred even earlier, but that perhaps these early domestic dogs went extinct (Ovodov et al., 2011; Germonpré et al., 2013). Numerous locations have been proposed for the origin of dogs, including Europe (Thalmann et al., 2013), the Middle East (Vonholdt et al., 2010), Africa (Boyko et al., 2009), Central Asia (Shannon et al., 2015), and Southeast Asia (Pang et al., 2009; Ding et al., 2012), but none have been widely accepted. The difficulty in pinpointing dogs' origin is compounded because most modern breeds were created only a few hundred years ago, in 18th-century Europe (Karlsson et al., 2007; Larson et al., 2012; Wayne and VonHoldt, 2012). By using modern dogs, it may only be possible to track dog demographic history to the most recent population replacement or breed formation event, not the advent of dog domestication (Sacks et al., 2013). More

recently, interest has shifted towards analyzing ancient dog remains, to bypass concerns regarding modern dog demography, and this has shed new light on dog domestication (Thalmann et al., 2013; Freedman et al., 2014; Frantz et al., 2016). For example, it was long thought that a single geographic origin of dogs was likely, considering the genetic homogeneity of modern dogs worldwide (Pang et al., 2009; Ardalan et al., 2011; Freedman et al., 2014), but more recently it has been suggested that there were two origins of domestication, and that one population replaced the other long before the creation of modern dog breeds (Frantz et al., 2016).

Dogs in the Americas

Dogs migrated with humans to the Americas across the Bering Land Bridge (Leonard et al., 2002), and were not domesticated from North American wolves. Some admixture with North American wolves has been inferred, but it seems to have occurred only rarely, and primarily in the Arctic (Koop et al., 2000). Dogs were widespread across North America by at least 9000 ybp, and likely entered South America much later, closer to 1500 ybp (Morey and Wiant, 1992; Schwartz, 1997; Yohe and Pavesic, 2000). This timing suggests that dogs may not have arrived with humans during the initial 16 kybp peopling of the Americas (Witt et al., 2015). Dogs were utilized by many Native American peoples in different ways: as a food source, as aids for hunting and fishing, and as load-bearers, guards, and pets (Schwartz, 1997). The usage of dogs in the Americas also changed over time; for example, dogs in the Midwest transitioned from being ceremonially buried during the Woodland period, from 1000-3000 ybp (Cantwell, 1980), to being used as a food source in the Mississippian period, starting at 1000 ybp (Borgic and Galloy, 2004). The largest numbers of dog burials can be found in the Southeastern United States dating to the Archaic period, approximately 3000-9000 ybp (Morey, 2006), and in the Midwest dating to the Woodland period (Cantwell, 1980; Lapham, 2010). However, dog burials have been found across North America and Mexico, as well as in South America in small numbers (Morey, 2006). While dogs had varied roles in different time periods and geographic regions, they were an important part of humans' lives in the Americas, and this

places them as likely good proxies to use to examine human history in the Americas.

Using Biological Proxies

A biological proxy, or bioproxy, is an organism that can be used to study a different taxon, if the latter is unavailable for study or if it yields limited information. The study of human demographic history is of interest to many researchers, as well as the public, but the specifics of the routes humans took or the different populations that interacted are largely unknown today. To try and clarify these gaps in understanding, a variety of species have been studied to learn more about human history. The largest case study for this is the peopling of Oceania (Matisoo-Smith and Robins, 2004; Larson et al., 2007; Storey et al., 2012; Thomson et al., 2014). Several species, including chickens, pigs, and rats, all were brought with humans as they moved from island to island, and the demographic history of these species has been studied to help understand how humans peopled Oceania. As another example, mice spread all over the world as stowaways on ships, and by studying their mitochondrial diversity, one can retrace early human voyages, including the travels of the Vikings and Phoenicians (Jones et al., 2013). In other parts of the world, parasites (Ascunce et al., 2013) and bacteria (Kersulyte et al., 2010; Breurec et al., 2013) have also been used to examine human demographic history as well.

Dogs have been used as proxies for humans in terms of adaptation, migration, and diet. In some cases, dogs and humans adapted to new environments in similar ways. For example, Tibetan mastiffs showed genetic changes to survive in high-altitude environments that are paralogous to human high-altitude adaptations (Li et al., 2014). Additionally, dogs have shown adaptation to a high-starch diet through an increase in copy number of salivary amylase, as have humans (Axelsson et al., 2013). Dogs that historically derive from regions of the world where domestic crops were utilized have a higher copy number of the amylase gene than dogs that do not (Freedman et al., 2014). This difference in copy number mirrors that of human populations with high and low starch diets (Perry et al., 2007). Dog populations have also been examined for their demographic history, to relate their history to that of humans. Dogs

have had a close relationship with humans for millennia, and when humans migrated, in many cases they would migrate with their dogs (Leonard et al., 2002; Ardalan et al., 2015). Given this shared population history, dogs and humans should show similar patterns both of genetic divergence from source populations, and of shared genetic variants between related populations. If both dog and human DNA are available from a region or an archaeological site, the two histories can be compared. If there are no human remains available for study, the dogs can instead be used to infer human movements. In the arctic, mitochondrial haplotype continuity in dogs for 700 years in both Alaska and Greenland signifies population continuity in the area (Brown et al., 2013). Additionally, the demography of the New Guinea singing dog and the dingo has been used to study the migration of Polynesians (Sacks et al., 2013; Greig et al., 2015). Dogs may not be a perfect proxy for humans in all cases. For example, dogs have been used as trade commodities (White et al., 2001; Rick et al., 2008), and so movement of dogs may not necessarily imply movement of humans. In the Americas, some peoples did not actively raise or keep dogs, and only interacted with them as puppies, with adult dogs being feral (Schwartz, 1997). In cases like this, the dogs' demographic history would be considered largely separate from that of humans, as the movement of human populations would not affect feral dogs.

Isotopically, dogs have been used as dietary proxies for humans as well. For example, dogs have been used to examine the transition from hunting and gathering to farming in Denmark (Noe-Nygaard, 1988). Mesolithic populations along the coast had average $\delta^{13}C$ values of -12 to -15‰, consistent with a diet of primarily marine resources, while Neolithic populations had $\delta^{13}C$ values of around -20‰, which is consistent with consuming more terrestrial plants due to agriculture. Dogs at these sites show the same shifts in stable isotope values. Dogs have also been used document maize consumption in Mississippian (Hogue, 2003; Allitt et al., 2009) and Mayan dogs (White et al., 2001). At archaeological sites in the Southern (Hogue, 2003) and the Northeastern United States (Allitt et al., 2009), dog remains have been used in lieu of human remains to determine the extent of maize consumption during the Mississippian

period. In both cases, the dog stable isotope values were similar to human stable isotope values from nearby archaeological sites from the same period, so it was possible to infer human diet at those sites. On the Channel Islands in California, stable isotope analysis of collagen from human, dog, and island fox (*Urocyon littoralis*) remains demonstrated that the humans and dogs had similar diets, while the island foxes had significantly lower $\delta^{13}C$ and $\delta^{15}N$ values (Rick et al., 2011). There is some debate about how closely isotopic measures in dogs relate to those in humans from related sites, but in general they show a high correlation (Guiry, 2012).

Human remains can be considered sacred by their descendants. For some, any research that involves destructive analysis of remains is unacceptable, because ancestral remains must stay whole. In the United States, there has been a long history of mistrust of scientists by Native Americans (Watkins, 2004; Bruning, 2006; Garrison, 2012; Bardill, 2014). Much of this mistrust stems from a long history of exploitation, mistreatment, and forcible removal of cultural identity by Europeans (Duran et al., 1998; Bowekaty and Davis, 2003; Wolfe, 2006). Many early anthropologists had a Western-centric perspective, and used anthropometric and genetic data to support eugenics and the idea of races as a biological construct, with some races considered to be "superior" to others (Provine and Smith, 1986; Bruce, 2000). This is also partially due to researchers taking samples and studying them in ways that peoples did not consent to (Garrison, 2012; Kowal, 2013; Bardill, 2014). The most prominent case of this misuse of DNA samples involves the Havasupai tribe, who had donated DNA samples for a study on diabetes (Garrison, 2012). Those samples were also used for other studies, for purposes that the Havasupai people had not given consent for, such as research on schizophrenia and inbreeding, both of which are taboo for the Havasupai. They sued Arizona State University and the Arizona Board of Research, resulting in a settlement, and this case had far-reaching consequences, both for scientists and for other Native American groups. Many Native American groups became even more hesitant to participate in genetic studies as a result of the Havasupai case, and some communities, including the Navajo Nation, have a

moratorium on all genetic research, although their reasons for the moratorium were different (Garrison, 2012).

Additionally, when ancient human remains are uncovered, there is often a disagreement between scientists, who wish to analyze the individual, and possible descendant peoples, who wish to simply repatriate and rebury them (Watkins, 2004). The most prominent recent example is Ancient One (also known as Kennewick Man), who was first discovered in 1996 but not repatriated until over 20 years later, in 2017 (Bruning, 2006). Despite the fact that Native American groups wanted to repatriate the Ancient One, multiple genetic and archaeological analyses were conducted (Bruning, 2006; Watson, 2015), culminating in a complete genome sequence (Rasmussen et al., 2015), and only then was repatriation of the Ancient One completed. More recently, some researchers have developed research projects with living communities, who wish to learn more about their history (Cui et al., 2013; Lindo et al., 2016). These projects involve consistent consulting between scientists, including archaeologists and geneticists, and the Native community, to ensure that the research trajectory is something both parties agree with. This is considered to be a more ethical way of studying the history of humans in the Americas, compared to past studies of Native Americans in which scientists took samples from a community and never returned to discuss the project or its findings. Using dogs as a proxy to study humans is another way to continue this research while respecting the wishes of communities who do not want their ancestors to be destructively analyzed. Dogs are considered to be sacred to some Native groups, including the Pueblo (Schwartz, 1997), but in many cases the destructive analysis of ancient dogs is preferable to the destructive analysis of human remains.

Ancient DNA

Ancient DNA is defined as any DNA that has been degraded due to environmental exposure. Often, this refers to archaeological remains, but this also applies to forensic remains as well. As an organism decomposes, cells burst open and the DNA inside them is exposed to water, heat, and sunlight,

as well as enzymes from decomposing bacteria and the organism itself (Hofreiter et al., 2001; Gilbert et al., 2003; Pääbo et al., 2004; J Dabney et al., 2013). DNA degrades rapidly after an organism dies - it has been estimated that 100 base pair (bp) fragments of DNA have a half-life of 150 years at a temperature of 25° C (Allentoft et al., 2012). Certain environmental conditions, such as cold temperatures, dryness, and protection from UV radiation, can extend the survival time of DNA. For example, using the same DNA decay model as previously mentioned, but at a temperature of 5° C, a 100 bp fragment of DNA has a much longer half-life of 6000 years (Allentoft et al., 2012). DNA has been successfully recovered from organisms at old as 600,000 ybp that were found in permafrost (Jesse Dabney et al., 2013; Orlando et al., 2013; Schubert et al., 2014; Skoglund et al., 2015). On the other hand, exposure to heat and moisture can cause DNA to degrade much faster. In North America, with a much more temperate climate, ancient DNA has been recovered from a few humans and dogs that are older than 9000 years (Kemp et al., 2007; Jenkins et al., 2012; Thalmann et al., 2013; Chatters et al., 2014; Rasmussen et al., 2014, 2015; Tackney et al., 2015; Lindo et al., 2017).

Working with ancient DNA presents unique challenges that require the use of special techniques to overcome. Ancient DNA accumulates damage from UV radiation, hydrolysis and oxidation (Gilbert et al., 2003; Willerslev and Cooper, 2005; J Dabney et al., 2013). This can result in strand breaks, causing the DNA to fragment into small segments. These segments are often shorter than 150 base pairs (bp) (Pääbo, 1989), so many primer pairs developed for amplifying modern DNA will not amplify ancient DNA. Sequencing an ancient mitochondrial genome (which is 16,000 bp long) using Sanger sequencing would require dozens of primer pairs, and so next-generation sequencing techniques, which can sequence many different fragments simultaneously, are more commonly used for ancient DNA sequencing. Damage can also cause depurination, in which a nucleotide base is cleaved from the DNA strand completely, making that segment of the DNA strand more prone to fragmentation (Gilbert, 2006). In some cases, the base pairs can even be directly altered. The most common base pair alteration is cytosine deamination, which

turns the cytosine into uracil (Hofreiter et al., 2001; Gilbert et al., 2003; Pääbo et al., 2004). This changes the DNA sequence, and can cause misinterpretation of sequence data, because the uracil will be replicated as a thymine. This effect of damage can often be mitigated through next generation sequencing, in which individual sequence reads can be compared, to help distinguish between the original sequence and the changes caused by damage.

In addition to DNA damage, contamination with modern DNA can also be problematic. This DNA can come from the archaeologists excavating the sample, the researchers working with it, or even from contaminated lab reagents. Modern DNA lacks the strand breaks and damage found in ancient DNA, and so it is much more likely to be amplified than the ancient DNA (Malmström et al., 2005). "Clean" excavations, in which samples that will be used in ancient DNA analysis are handled with gloves throughout the excavation process and are not washed, which is a frequent source of contamination, can help prevent the introduction of modern DNA to the sample, but are rare (Pruvost et al., 2007; Adler et al., 2011; Meyer et al., 2016).

Many methods and guidelines have been developed for working with ancient DNA, to minimize contamination and maximize DNA yield (Cooper and Poinar, 2000; Kaestle and Horsburgh, 2002; Adler et al., 2011; Barta et al., 2014). To limit contamination, a laboratory dedicated to extracting DNA from ancient individuals must be physically removed from the lab where modern DNA is extracted. All researchers working with ancient DNA wear protective full-body clothing to avoid contamination, and all lab equipment is wiped down with bleach and treated with UV light to destroy or crosslink any DNA that remains. DNA recovery methods that favor small fragment sizes have been developed to maximize extraction efficiency, including the use of PCR purification kits (Yang et al., 1998) or silica solutions (Allentoft et al., 2015). Once DNA has been sequenced, it is common to extract DNA from the same individual multiple times to confirm that the sequence is accurate, and certain properties of ancient DNA (damaged ends and short fragments) can be used to confirm that the DNA is ancient (Jónsson et al., 2013).

Additionally, the individuals working with ancient DNA often have their own DNA sequenced, to compare to the ancient samples and make sure contamination is not a concern. With these safeguards in place, ancient DNA can be reliably recovered and authenticated.

Thesis Outline

This thesis includes three data chapters, each of which is formatted as a separate paper. The first chapter, which was published in the Journal of Human Evolution in 2015 (Witt et al., 2014), examines the hypervariable region (HVR) of mitochondrial DNA in 42 ancient dogs from three archaeological sites, which is compared to nearly all published ancient dog mitochondrial DNA sequences. Populations were compared in terms of genetic diversity, and shared or closely related haplotypes between populations were identified. This study demonstrated that there was a single haplotype that was common across North America, and that different populations had different levels of diversity, suggesting that dogs may have been deliberately bred in some regions of the Americas, including the Midwest, or that they came from small founding populations. Additionally, some Arctic dogs had mitochondrial DNA sequences that were most similar to that of wolves, suggesting that there may have been some dog-wolf admixture in the Arctic. Demographic modeling of ancient dogs in the Americas suggested that dogs may have arrived in the Americas as recently as 10,000 ybp.

The second chapter takes a similar approach to the first, but reports on complete mitogenome sequences, produced with high-throughput sequencing techniques. In this study, a total of 69 ancient dogs from 19 archaeological sites were sequenced, and compared to three published mitochondrial genomes from ancient dogs in the Americas (Thalmann et al. 2013). This study also assessed population genetic diversity levels, and compared the populations to one another and to modern dogs and wolves to find shared or closely related haplotypes. This research found that sequencing the mitogenome yields a much higher resolution view of dog population diversity. Contrary to previous research, ancient dogs and modern dogs do not share mitochondrial haplotypes, and this suggests that there was a large loss in dog

diversity following European contact. Ancient American dogs' mitogenomes are most closely related to the mitogenomes of wolves from Siberia and Switzerland, supporting the hypothesis that dogs migrated with humans to the Americas, rather than being domesticated there separately. Similar to the mitochondrial genetic structuring found in Native American mitogenomes, dog mitogenomes form two major clades, each with coalescence dates of 13,000 to 17,0000 ybp. Additionally, dog populations show affinity between Midwest and Southeast populations, as well as populations along the Pacific Coast. Increases in dog genetic diversity over time in the Midwest were found to be coincident with the transition between Late Woodland (3000-1000 ybp) and Mississippian periods (1000-600 ybp), marked by a shift from small-scale horticulture to large-scale maize agriculture and population concentration in city centers. Finally, demographic modeling of dog diversity over time showed that dogs migrated to the Americas between 17,000 ybp and 12,000 ybp, and that the dog population may have begun to decline around 2000 ybp, well before Europeans arrived to the Americas.

The third chapter is focused on an archaeological site in Southern Illinois that was occupied through the Woodland-Mississippian transition, known as Janey B. Goode (approximately 1100-800 AD). The Mississippians were maize agriculturalists, but the timing of the arrival of maize is uncertain. There are some sites in the Southeastern United States with maize present as early as the Middle Woodland period, over 3000 ybp (Fearn and Liu, 1995), but the earliest evidence for maize in Southern Illinois dates to 900 AD (Vanderwarker et al., 2013; Simon, 2017). Late Woodland populations in Southern Illinois grew a number of crops including squash and sumpweed (Smith, 1989; Simon and Lopinot, 2006; Simon, 2010), and over time maize consumption slowly increased, making up 40-50% of the diet during the Early Mississippian period (1000-1100 AD) and increasing to as much as 80% of the diet during the Late Mississippian period (1400-1600 AD) (Hedman et al., 2002; Emerson et al., 2005; Yerkes, 2011). What is known about maize intensification in the region is primarily identified from the $\delta^{13}C$ of human bones (Ambrose, 1987), but the human remains in the region are from the Mississippian period, when maize

was already established. Dog remains from the site date to the time of that transition, and so they are used as a dietary proxy for humans to assess when maize consumption began to increase. This is accomplished through stable isotope analysis of dog bones and teeth, which shows general dietary trends, as well as shotgun sequencing of dog coprolites to examine specific dietary components. This research shows an increase of $\delta^{13}C$ between the Woodland and Mississippian periods, signifying an increase in maize consumption over time. The $\delta^{15}N$ value is low and the $\delta^{18}O$ value is high, suggesting that plants were a large proportion of the dogs' diet across the Late Woodland and Mississippian periods. DNA sequences from the coprolites show that the dogs ate maize, and they were also eating squash, nightshade, tobacco, herons, and multiple species of fish. The dogs' stable isotope values fit with contemporaneous human populations from the Midwest, suggesting that the dogs and humans at Janey B. Goode ate very similar diets. *Toxocara canis*, a parasitic nematode, was also identified in multiple dog coprolites, which suggests that this likely affected the health of both humans and dogs during the Late Woodland and Mississippian periods. By using the dogs as a dietary proxy for humans, we determined that humans during the Mississippian period likely ate large amounts of maize, along with squash, tobacco and nightshade, as well as herons and multiple species of fish.

References

Adler, C.J., Haak, W., Donlon, D., Cooper, A., 2011. Survival and recovery of DNA from ancient teeth and bones. Journal of Archaeological Science. 38, 956–964.

Allentoft, M.E., Collins, M., Harker, D., Haile, J., Oskam, C.L., Hale, M.L., Campos, P.F., Samaniego, J. a, Gilbert, M.T.P., Willerslev, E., Zhang, G., Scofield, R.P., Holdaway, R.N., Bunce, M., 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proceedings of the Royal Society B: Biological Sciences. 279, 4724–33.

Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., Malaspinas, A., Margaryan, A., Higham, T., Chivall, D., Lynnerup, N., Harvig, L., Baron, J., Casa, P. Della, Dąbrowski, P., Duffy, P.R., Ebel, A. V., Epimakhov, A., Frei, K., Furmanek, M., Gralak, T., Gromov, A., Gronkiewicz, S., Grupe, G., Hajdu, T., Jarysz, R., Khartanovich, V., Khokhlov, A., Kiss, V., Kolář, J., Kriiska, A., Lasak, I., Longhi, C., McGlynn, G., Merkevicius, A., Merkyte, I., Metspalu, M., Mkrtchyan, R., Moiseyev, V., Paja, L., Pálfi, G., Pokutta, D., Pospieszny, Ł., Price, T.D., Saag, L., Sablin, M., Shishlina, N., Smrčka, V., Soenov, V.I., Szeverényi, V., Tóth, G., Trifanova, S. V., Varul, L., Vicze, M., Yepiskoposyan, L., Zhitenev, V., Orlando, L., Sicheritz-Pontén, T., Brunak, S., Nielsen, R., Kristiansen, K., Willerslev, E., 2015. Population genomics of Bronze Age Eurasia. Nature. 522, 167–172.

Allitt, S., Stewart, R.M., Messner, T., 2009. The Utility of Dog Bone (*Canis Familiaris*) in Stable Isotope Studies for Investigating the Presence of Prehistoric Maize (*Zea mays ssp. mays*): A Preliminary Study. North American Archaeologist. 29, 343–367.

Ambrose, S., 1987. Chemical and isotopic techniques of diet reconstruction in eastern North America. In: Emergent Horticultural Economies of the Eastern Woodlands. pp. 87-107.

Ardalan, A., Kluetsch, C.F.C., Zhang, A.-B., Erdogan, M., Uhlén, M., Houshmand, M., Tepeli, C., Ashtiani, S.R.M., Savolainen, P., 2011. Comprehensive study of mtDNA among Southwest Asian dogs contradicts independent domestication of wolf, but implies dog-wolf hybridization. Ecology and Evolution. 1, 373–85.

Ardalan, A., Oskarsson, M.C.R., van Asch, B., Rabakonandriania, E., Savolainen, P., 2015. African origin for Madagascan dogs revealed by mtDNA analysis. Royal Society Open Science. 2, 140552–140552.

Ascunce, M.S., Fane, J., Kassu, G., Toloza, A.C., Picollo, M.I., González-Oliver, A., Reed, D.L., 2013. Mitochondrial diversity in human head louse populations across the Americas. American Journal of Physical Anthropology. 152, 118–29.

Axelsson, E., Ratnakumar, A., Arendt, M.-L., Maqbool, K., Webster, M.T., Perloski, M., Liberg, O., Arnemo, J.M., Hedhammar, A., Lindblad-Toh, K., 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature. 495, 360-364.

Bardill, J., 2014. Native American DNA: Ethical, Legal, and Social Implications of an Evolving Concept. Annual Reviews in Anthropology. 43, 155-166.

Barta, J.L., Monroe, C., Teisberg, J.E., Winters, M., Flanigan, K., Kemp, B.M., 2014. One of the key characteristics of ancient DNA, low copy number, may be a product of its extraction. Journal of

Archaeological Science. 46, 281–289.

Borgic, Q.L., Galloy, J.M., 2004. Domesticated Dog Remains from the Janey B. Goode Site. In: Joint Meetings of the Southeastern Archaeological Conference and the Midwestern Archaeological Conference, St. Louis, Missouri, October (Vol. 2).

Bowekaty, M., and Davis, D., 2003. Cultural issues in genetic research with American Indian and Alaskan Native people. IRB Ethics and Human Research. 4, 12-15.

Boyko, A.R., Boyko, R.H., Boyko, C.M., Parker, H.G., Castelhano, M., Corey, L., Degenhardt, J.D., Auton, A., Hedimbi, M., Kityo, R., Ostrander, E. A., Schoenebeck, J., Todhunter, R.J., Jones, P., Bustamante, C.D., 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. Proceedings of the National Academy of Sciences of the United States of America. 106, 13903–8.

Breurec, S., Raymond, J., Thiberge, J.-M., Hem, S., Monchy, D., Seck, A., Dehoux, P., Garin, B., Dauga, C., 2013. Impact of human migrations on diversity of Helicobacter pylori in Cambodia and New Caledonia. Helicobacter. 18, 249–61.

Brown, S.K., Darwent, C.M., Sacks, B.N., 2013. Ancient DNA evidence for genetic continuity in arctic dogs. Journal of Archaeological Science. 40, 1279–1288.

Bruce, B., 2000. Working toward an Ethical Anthropology: Where Do We Go From Here? Annals of the New York Academy of Sciences. 925, 187-195.

Bruning, S.B., 2006. Complex Legal Legacies : The Native American Graves Protection and Repatriation Act , Scientific Study , and Kennewick Man. American Antiquity. 71, 501–521.

Cantwell, A., 1980. Middle Woodland dog ceremonialism in Illinois. The Wisconsin Archeologist. 61, 480–496.

Chatters, J.C., Kennett, D.J., Asmerom, Y., Kemp, B.M., Polyak, V., Blank, A.N., Beddows, P. A., Reinhardt, E., Arroyo-Cabrales, J., Bolnick, D. A., Malhi, R.S., Culleton, B.J., Erreguerena, P.L., Rissolo, D., Morell-Hart, S., Stafford, T.W., 2014. Late Pleistocene human skeleton and mtDNA link Paleoamericans and modern Native Americans. Science. 344, 750–4.

Clutton-Brock, J., 1995. Origins of the dog: domestication and early history. In: Serpell, J. (Ed.), The Domestic Dog: Its Evolution, Behaviour and Interactions with People. Cambridge University Press, Cambridge, pp. 7–20.

Cooper, A., Poinar, H., 2000. Ancient DNA: do it right or not at all. Science. 289, 10–11.

Cui, Y., Lindo, J., Hughes, C.E., Johnson, J.W., Hernandez, A.G., Kemp, B.M., Ma, J., Cunningham, R., Petzelt, B., Mitchell, J., Archer, D., Cybulski, J.S., Malhi, R.S., 2013. Ancient DNA analysis of mid-holocene individuals from the Northwest Coast of North America reveals different evolutionary paths for mitogenomes. PloS one. 8, e66948.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., Meyer, M., 2013. Complete mitochondrial genome sequence of a Middle

Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences of the United States of America. 110, 15758–63.

Dabney, J., Meyer, M., Paabo, S., 2013. Ancient DNA Damage. Cold Spring Harbor Perspectives in Biology. 5, a012567–a012567.

Ding, Z.-L., Oskarsson, M., Ardalan, A., Angleby, H., Dahlgren, L.-G., Tepeli, C., Kirkness, E., Savolainen, P., Zhang, Y.-P., 2012. Origins of domestic dog in southern East Asia is supported by analysis of Y-chromosome DNA. Heredity. 108, 507–14.

Duran, B., Duran, E., and Brave Heart, M. Y. H, 1998. Native Americans and the trauma of history. In: Studying native America: Problems and prospects. pp. 60-76.

Emerson, T.E., Hedman, K.M., Simon, M.L., 2005. Marginal horticulturalists or maize agriculturalists? Archaeobotanical, paleopathological, and isotopic evidence relating to Langford tradition maize consumption. Midcontinental Journal of Archaeology. 30, 67–118.

Fearn, M.L., Liu, K.-B., 1995. Maize Pollen of 3500 B.P. From Southern Alabama. American Antiquity. 60, 109–117.

Frantz, L.A.F., Mullin, V.E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., Linderholm, A., Mattiangeli, V., Teasdale, M.D., Dimopoulos, E.A., Tresset, A., Duffraisse, M., McCormick, F., Bartosiewicz, L., Gal, E., Nyerges, E.A., Sablin, M. V., Brehard, S., Mashkour, M., Balasescu, A., Gillet, B., Hughes, S., Chassaing, O., Hitte, C., Vigne, J.-D., Dobney, K., Hanni, C., Bradley, D.G., Larson, G., 2016. Genomic and archaeological evidence suggests a dual origin of domestic dogs. Science. 352, 1228–1231.

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E. a., Marques-Bonet, T., Wayne, R.K., Novembre, J., 2014. Genome Sequencing Highlights the Dynamic Early History of Dogs. PLoS Genetics. 10, e1004016.

Garrison, N. a., 2012. Genomic Justice for Native Americans: Impact of the Havasupai Case on Genetic Research. Science, Technology & Human Values. 38, 201–223.

Germonpré, M., Sablin, M. V., Després, V., Hofreiter, M., Lázničková-Galetová, M., Stevens, R.E., Stiller, M., 2013. Palaeolithic dogs and the early domestication of the wolf: a reply to the comments of Crockford and Kuzmin (2012). Journal of Archaeological Science. 40, 786–792.

Gilbert, M.T.P., 2006. Postmortem damage of mitochondrial DNA. In: Human Mitochondrial DNA and the Evolution of of *Homo sapiens*. pp. 91-115.

Gilbert, M.T.P., Hansen, A.J., Willerslev, E., Rudbeck, L., Barnes, I., Lynnerup, N., Cooper, A., 2003. Characterization of genetic miscoding lesions caused by postmortem damage. American Journal of Human Genetics. 72, 48–61.

Greig, K., Boocock, J., Prost, S., Horsburgh, K.A., Jacomb, C., Walter, R., Matisoo-Smith, E., 2015.

Complete Mitochondrial Genomes of New Zealand's First Dogs. Plos ONE. 10, e0138536.

Guiry, E.J., 2012. Dogs as Analogs in Stable Isotope-Based Human Paleodietary Reconstructions: A Review and Considerations for Future Use. Journal of Archaeological Method and Theory. 19, 351–376.

Hedman, K., Hargrave, E. A., Ambrose, S.H., 2002. Late Mississippian Diet in the American Bottom: Stable Isotope Analyses of Bone Collagen and Apatite. Midcontinental Journal of Archaeology. 27, 237–271.

Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., Pääbo, S., 2001. Ancient DNA. Nature reviews Genetics. 2, 353–9.

Hogue, S., 2003. Corn dogs and hush puppies: diet and domestication at two protohistoric farmsteads in Oktibbeha County, Mississippi. Southeastern Archaeology. 22, 185–195.

Jenkins, D.L., Davis, L.G., Stafford, T.W., Campos, P.F., Hockett, B., Jones, G.T., Cummings, L.S., Yost, C., Connolly, T.J., Yohe II, R.M., Gibbons, S.C., Raghavan, M., Rasmussen, M., Paijmans, J.L.A., Hofreiter, M., Kemp, B.M., Barta, J.L., Monroe, C., Gilbert, M.T.P., Willerslev, E., 2012. Clovis Age Western Stemmed Projectile Points and Human Coprolites at the Paisley Caves. Science. 337, 223–228.

Jones, E.P., Eager, H.M., Gabriel, S.I., Jóhannesdóttir, F., Searle, J.B., 2013. Genetic tracking of mice and other bioproxies to infer human history. Trends in Genetics. 29, 298–308.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics. 29, 1682–4.

Kaestle, F. A., Horsburgh, K.A., 2002. Ancient DNA in anthropology: Methods, applications, and ethics. American Journal of Physical Anthropology. 119, 92–130.

Karlsson, E.K., Baranowska, I., Wade, C.M., Salmon Hillbertz, N.H.C., Zody, M.C., Anderson, N., Biagi, T.M., Patterson, N., Pielberg, G.R., Kulbokas, E.J., Comstock, K.E., Keller, E.T., Mesirov, J.P., von Euler, H., Kämpe, O., Hedhammar, A., Lander, E.S., Andersson, G., Andersson, L., Lindblad-Toh, K., 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. Nature Genetics. 39, 1321–1328.

Kemp, B., Malhi, R., McDonough, J., Bolnick, D.A., Eshleman, J.A., Rickards, O., Martinez-Labarga, C., Johnson, J.R., Lorenz, J.G., Dixon, E.J., Fifield, T.E., Heaton, T.H., Worl, R., Smith, D.G., 2007. Genetic Analysis of Early Holocene Skeletal Remains From Alaska and its Implications for the Settlement of the Americas. American Journal of Physical Anthropology. 132, 605–621.

Kersulyte, D., Kalia, A., Gilman, R.H., Mendez, M., Herrera, P., Cabrera, L., Velapatiño, B., Balqui, J., Paredes Puente de la Vega, F., Rodriguez Ulloa, C. A, Cok, J., Hooper, C.C., Dailide, G., Tamma, S., Berg, D.E., 2010. Helicobacter pylori from Peruvian amerindians: traces of human migrations in strains from remote Amazon, and genome sequence of an Amerind strain. PloS ONE. 5, e15076.

Koop, B.F., Burbidge, M., Byun, A., Rink, U., Crockford, S., 2000. Ancient DNA evidence of a separate origin for North American indigenous dogs. In: Crockford, S.J. (Ed.), Dogs Through Time: An

Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. Archaeopress.

Kowal, E., 2013. Orphan DNA: Indigenous samples, ethical biovalue and postcolonial science. Social Studies of Science. 43, 577–597.

Lapham, H.A., 2010. A Baumer Phase Dog Burial from the Kincaid Site in Southern Illinois. Illinois Archaeology. 22, 437–463.

Larson, G., Cucchi, T., Fujita, M., Matisoo-Smith, E., Robins, J., Anderson, A., Rolett, B., Spriggs, M., Dolman, G., Kim, T.-H., Thuy, N.T.D., Randi, E., Doherty, M., Due, R.A., Bollt, R., Djubiantono, T., Griffin, B., Intoh, M., Keane, E., Kirch, P., Li, K.-T., Morwood, M., Pedriña, L.M., Piper, P.J., Rabett, R.J., Shooter, P., Van den Bergh, G., West, E., Wickler, S., Yuan, J., Cooper, A., Dobney, K., 2007. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proceedings of the National Academy of Sciences of the United States of America. 104, 4834–9.

Larson, G., Karlsson, E.K., Perri, A., Webster, M.T., Ho, S.Y.W., Peters, J., Stahl, P.W., Piper, P.J., Lingaas, F., Fredholm, M., Comstock, K.E., Modiano, J.F., Schelling, C., Agoulnik, A.I., Leegwater, P. a, Dobney, K., Vigne, J.-D., Vilà, C., Andersson, L., Lindblad-Toh, K., 2012. Rethinking dog domestication by integrating genetics, archeology, and biogeography. Proceedings of the National Academy of Sciences of the United States of America. 109, 8878–83.

Leonard, J. A., Wayne, R.K., Wheeler, J., Valadez, R., Guillén, S., Vilà, C., 2002. Ancient DNA evidence for Old World origin of New World dogs. Science. 298, 1613–6.

Li, Y., Wu, D.-D., Boyko, A.R., Wang, G.-D., Wu, S.-F., Irwin, D.M., Zhang, Y.-P., 2014. Population Variation Revealed High-Altitude Adaptation of Tibetan Mastiffs. Molecular Biology and Evolution. 31, 1200–1205.

Lindo, J., Huerta-Sanchez, E., Nakagome, S., Rasmussen, M., Petzelt, B., Mitchell, J., Cybulski, J.S., Willerslev, E., DeGiorgio, M., Malhi, R.S., 2016. A time transect of exomes from a Native American population before and after European contact. Nature Communications. 7, 13175.

Lindo, J., Achilli, A., Perego, U.A., Archer, D., Valdiosera, C., Petzelt, B., Mitchell, J., Worl, R., Dixon, E.J., Fifield, T.E., Rasmussen, M., Willerslev, E., Cybulski, J.S., Kemp, B.M., DeGiorgio, M., Malhi, R.S., 2017. Ancient individuals from the North American Northwest Coast reveal 10,000 years of regional genetic continuity. Proceedings of the National Academy of Sciences. 114, 4093-4098.

Malmström, H., Storå, J., Dalén, L., Holmlund, G., Götherström, A., 2005. Extensive human DNA contamination in extracts from ancient dog bones and teeth. Molecular Biology and Evolution. 22, 2040–7.

Matisoo-Smith, E., Robins, J.H., 2004. Origins and dispersals of Pacific peoples: evidence from mtDNA phylogenies of the Pacific rat. Proceedings of the National Academy of Sciences of the United States of America. 101, 9167–72.

Meyer, M., Arsuaga, J.-L., de Filippo, C., Nagel, S., Aximu-Petri, A., Nickel, B., Martínez, I., Gracia, A., de Castro, J.M.B., Carbonell, E., Viola, B., Kelso, J., Prüfer, K., Pääbo, S., 2016. Nuclear DNA sequences

from the Middle Pleistocene Sima de los Huesos hominins. Nature. 531, 504-507.

Morey, D.F., 2006. Burying key evidence : the social bond between dogs and people. Journal of Archaeological Science. 33, 158–175.

Morey, D.F., Wiant, M.D., 1992. Early Holocene Domestic Dog Burials from the North American Midwest. Current Anthropology. 33, 224–230.

Noe-Nygaard, N., 1988. δ 13 C-values of dog bones reveal the nature of changes in man's food resources at the mesolithic-neolithic transition, Denmark. Chemical Geology: Isotope Geoscience section. 13, 87–96.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M. V, Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K. a S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J. V, Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J., Willerslev, E., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature. 499, 74–8.

Ovodov, N.D., Crockford, S.J., Kuzmin, Y. V, Higham, T.F.G., Hodgins, G.W.L., van der Plicht, J., 2011. A 33,000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the Last Glacial Maximum. PloS ONE. 6, e22821.

Pääbo, S., 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proceedings of the National Academy of Sciences of the United States of America. 86, 1939–43.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., 2004. Genetic analyses from ancient DNA. Annual Review of Genetics. 38, 645–79.

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A., Luo, L.-Y., Angleby, H., Ardalan, A., Ekström, C., Sköllermo, A., Lundeberg, J., Matsumura, S., Leitner, T., Zhang, Y.-P., Savolainen, P., 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. Molecular Biology and Evolution. 26, 2849–64.

Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F. A., Mountain, J.L., Misra, R., Carter, N.P., Lee, C., Stone, A.C., 2007. Diet and the evolution of human amylase gene copy number variation. Nature Genetics. 39, 1256–60.

Provine, W.B., and Russell, E.S., 1986. Geneticists and Race. American Zoologist. 26, 857-887.

Pruvost, M., Schwarz, R., Correia, V.B., Champlot, S., Braguier, S., Morel, N., Fernandez-Jalvo, Y., Grange, T., Geigl, E.-M., 2007. Freshly excavated fossil bones are best for amplification of ancient DNA. Proceedings of the National Academy of Sciences of the United States of America. 104, 739–44.

Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Rasmussen, S.,

Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir, V., Yadav, R., Malaspinas, A.-S., White, S.S., Allentoft, M.E., Cornejo, O.E., Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M., Korneliussen, T.S., Meltzer, D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S., Brunak, S., Sicheritz-Ponten, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica, A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., Willerslev, E., 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature. 506, 225–9.

Rasmussen, M., Sikora, M., Albrechtsen, A., Korneliussen, T.S., Moreno-Mayar, J.V., Poznik, G.D., Zollikofer, C.P.E., Ponce de León, M.S., Allentoft, M.E., Moltke, I., Jónsson, H., Valdiosera, C., Malhi, R.S., Orlando, L., Bustamante, C.D., Stafford, T.W., Meltzer, D.J., Nielsen, R., Willerslev, E., 2015. The ancestry and affiliations of Kennewick Man. Nature. 523, 455-458.

Rick, T.C., Culleton, B.J., Smith, C.B., Johnson, J.R., Kennett, D.J., 2011. Stable isotope analysis of dog, fox, and human diets at a Late Holocene Chumash village (CA-SRI-2) on Santa Rosa Island, California. Journal of Archaeological Science. 38, 1385–1393.

Rick, T.C., Walker, P.L., Willis, L.M., Noah, A. C., Erlandson, J.M., Vellanoweth, R.L., Braje, T.J., Kennett, D.J., 2008. Dogs, humans and island ecosystems: the distribution, antiquity and ecology of domestic dogs (Canis familiaris) on California's Channel Islands, USA. The Holocene. 18, 1077–1087.

Sacks, B.N., Brown, S.K., Stephens, D., Pedersen, N.C., Wu, J.-T., Berry, O., 2013. Y chromosome analysis of dingoes and southeast asian village dogs suggests a neolithic continental expansion from southeast Asia followed by multiple austronesian dispersals. Molecular Biology and Evolution. 30, 1103–18.

Savolainen, P., Leitner, T., Wilton, A.N., Matisoo-Smith, E., Lundeberg, J., 2004. A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. Proceedings of the National Academy of Sciences of the United States of America. 101, 12387–90.

Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A., Albrechtsen, A., Dupanloup, I., Foucal, A., Petersen, B., Fumagalli, M., Raghavan, M., Seguin-Orlando, A., Korneliussen, T.S., Velazquez, A.M. V., Stenderup, J., Hoover, C. a., Rubin, C.-J., Alfarhan, A.H., Alquraishi, S. A., Al-Rasheid, K. a. S., MacHugh, D.E., Kalbfleisch, T., MacLeod, J.N., Rubin, E.M., Sicheritz-Ponten, T., Andersson, L., Hofreiter, M., Marques-Bonet, T., Gilbert, M.T.P., Nielsen, R., Excoffier, L., Willerslev, E., Shapiro, B., Orlando, L., 2014. Prehistoric genomes reveal the genetic foundation and cost of horse domestication. Proceedings of the National Academy of Sciences. 111, 201416991.

Schwartz, M., 1997. A History of Dogs in the Early Americas. Yale University Press, New Haven, CT.

Shannon, L.M., Boyko, R.H., Castelhano, M., Corey, E., Hayward, J.J., Mclean, C., White, M.E., Abi, M., Anita, B.A., Ikombe, N., Calero, J., Galov, A., 2015. Genetic structure in village dogs reveals a Central Asian domestication origin. Proceedings of the National Academy of Sciences of the United States of America. 112, 13639-13644.

Simon, M., 2010. Beyond Gathering : Early Plant Cultivation in Illinois. Illinois Antiquity. 45, 17–21.

Simon, M.L., 2017. Reevaluating the Evidence for Middle Woodland Maize From the Holding Site. American Antiquity. 82, 140–150.

Simon, M.L., Lopinot, K.E., 2006. PREHISTORIC PLANT USE IN THE AMERICAN BOTTOM : NEW THOUGHTS AND INTERPRETATIONS. Southeastern Archaeology. 25, 212–257.

Skoglund, P., Ersmark, E., Palkopoulou, E., Dalén, L., 2015. Ancient Wolf Genome Reveals an Early Divergence of Domestic Dog Ancestors and Admixture into High-Latitude Breeds. Current Biology. 11, 1515–1519.

Smith, B., 1989. Origins of agriculture in eastern North America. Science. 246, 1566–1571.

Storey, A. A., Athens, J.S., Bryant, D., Carson, M., Emery, K., deFrance, S., Higham, C., Huynen, L., Intoh, M., Jones, S., Kirch, P. V, Ladefoged, T., McCoy, P., Morales-Muñiz, A., Quiroz, D., Reitz, E., Robins, J., Walter, R., Matisoo-Smith, E., 2012. Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. PloS ONE. 7, e39171.

Tackney, J.C., Potter, B.A., Raff, J., Powers, M., Watkins, W.S., Warner, D., Reuther, J.D., Irish, J.D., O'Rourke, D.H., 2015. Two contemporaneous mitogenomes from terminal Pleistocene burials in eastern Beringia. Proceedings of the National Academy of Sciences. 112, 13833-13838.

Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpre, M.., Sablin, M.V., Lopez-Giraldez, F., Domingo-Roura, X., Napierala, H., Uerpmann, H.-P., Loponte, D.M., Acosta, A.., Giemsch, L., Schmitz, R.W., Worthington, B., Buikstra, J.E., Druzhkova, A., Graphodatsky, A.S., Ovodov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfli, K.-P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., Wayne, R.K., 2013. Complete Mitochondrial Genomes of Ancient Canids Suggest a European Origin of Domestic Dogs. Science. 342, 871–874.

Thomson, V.A., Lebrasseur, O., Austin, J.J., Hunt, T.L., Burney, D.A., Denham, T., Rawlence, N.J., Wood, J.R., Gongora, J., Girdland Flink, L., Linderholm, A., Dobney, K., Larson, G., Cooper, A., 2014. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences. 111, 4826–4831.

Vanderwarker, A.M., Wilson, G.D., Bardolph, D.N., 2013. Maize Adoption and Intensification in the Central Illinois River Valley: An Analysis of Archaeobotanical Data from the Late Woodland to Early Mississippian Periods (A.D. 600-1200). Southeastern Archaeology. 32, 147–168.

Vonholdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E., Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D. A., Auton, A., Reynolds, A., Bryc, K., Brisbin, A., Knowles, J.C., Mosher, D.S., Spady, T.C., Elkahloun, A., Geffen, E., Pilot, M., Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.-L., Zhang, Y.-P., Bustamante, C.D., Ostrander, E. a, Novembre, J., Wayne, R.K., 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature. 464, 898–902.

Watkins, J., 2004. Becoming American or Becoming Indian?: NAGPRA, Kennewick and Cultural Affiliation. Journal of Social Archaeology. 4, 60–80.

Watson, T., 2015. DNA shows 8 , 500-year-old Kennewick Man was Native American. USA Today. Published 06/18/2015.

Wayne, R.K., VonHoldt, B.M., 2012. Evolutionary genomics of dog domestication. Mammalian genome. 23, 3–18.

White, C.D., Pohl, M.E.D., Schwarcz, H.P., Longstaffe, F.J., 2001. Isotopic Evidence for Maya Patterns of Deer and Dog Use at Preclassic Colha. Journal of Archaeological Science. 28, 89–107.

Willerslev, E., Cooper, A., 2005. Ancient DNA. Proceedings of the Royal Society B: Biological Sciences. 272, 3–16.

Witt, K.E., Judd, K., Kitchen, A., Grier, C., Kohler, T. a., Ortman, S.G., Kemp, B.M., Malhi, R.S., 2015. DNA analysis of ancient dogs of the Americas: Identifying possible founding haplotypes and reconstructing population histories. Journal of Human Evolution. 79, 105–118.

Wolfe, P., 2006. Settler colonialism and the elimination of the native. Journal of Genocide Research. 8, 387-409.

Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., Saunders, S.R., 1998. Improved DNA Extraction From Ancient Bones Using Silica-Based Spin Columns. American Journal of Physical Anthropology. 543, 539–543.

Yerkes, R.W., 2011. Bone Chemistry, Body Parts, and Growth Marks: Evaluating Ohio Hopewell and Cahokia Mississippian Seasonality, Subsistence, Ritual, and Feasting. American Archaeology. 70, 241–265.

Yohe, R., Pavesic, M., 2000. Early archaic domestic dogs from Western Idaho, USA. In: Dogs Through Time: An Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. pp. 93–104.

CHAPTER TWO: DNA ANALYSIS OF ANCIENT DOGS OF THE AMERICAS: IDENTIFYING POSSIBLE FOUNDING HAPLOTYPES AND RECONSTRUCTING POPULATION HISTORIES.

Abstract

As dogs have traveled with humans to every continent, they can potentially serve as an excellent proxy when studying human migration history. Past genetic studies into the origins of ancient American dogs have used portions of the hypervariable region (HVR) of mitochondrial DNA (mtDNA) to indicate that prior to European contact, the dogs of Native Americans originated in Eurasia. In this study, we summarize past DNA studies of both humans and dogs to discuss their population histories in the Americas. We then sequence a portion of the mtDNA HVR of 42 pre-Columbian dogs from three sites located in Illinois, coastal British Columbia, and Colorado, and identify four novel dog mtDNA haplotypes. We next analyzed a dataset comprised of all available ancient dog sequences from the Americas to infer the pre-Columbian population history of dogs in the Americas. Interestingly, we found low levels of genetic diversity for some populations consistent with the possibility of deliberate breeding practices. Furthermore, we identified multiple putative founding haplotypes in addition to dog haplotypes that closely resemble those of wolves, suggesting admixture with North American wolves or perhaps a second domestication of canids in the Americas. Notably, initial effective population size estimates suggest at least 1,000 female dogs likely existed in the Americas at the time of the first known canid burial, and that population size increased gradually over time before stabilizing roughly 1,200 years before present.

Introduction

The domestic dog (*Canis lupus familiaris*) holds a unique place in the history of animal domestication, because this species was the first to be domesticated, and was also domesticated for a variety of purposes: as guards, hunting aids, and even as companions (Clutton-Brock, 1995). Dog remains dating to 10,000-14,000 years before present (ybp) have been discovered across Eurasia, and genetic studies suggest that dogs were domesticated from gray wolves between 11,000-20,000 years ago (Germonpré et al., 2009; Pang et al., 2009; Ding et al., 2012; Freedman et al., 2014). Recent analysis of an ancient Siberian canid with a morphology suggestive of a "transitional dog" and a mitochondrial DNA (mtDNA) haplotype found in contemporary dog populations suggests that domestication could have taken place in excess of 33,000 ybp (Druzhkova et al., 2013). The exact origin of domestic dogs is uncertain, though suggested geographic origins include the Middle East, Southeast Asia, Europe, and Africa (Boyko et al., 2009; Pang et al., 2009; Vonholdt et al., 2010; Ardalan et al., 2011; Ding et al., 2012; Thalmann et al., 2013). Most recently, however, results suggest that modern wolf populations diverged from one another at around the same time as dog domestication, and therefore modern populations cannot be used to determine where dogs were first domesticated (Freedman et al., 2014).

Dogs are found in a variety of archaeological contexts in the Americas that date as early as 10,500 ybp, with the first unequivocal dog burial dating to roughly 9,000 ybp (Morey and Wiant, 1992). Interestingly, genetic analysis of ancient dog mtDNA indicates that many of these dogs were domesticated from Eurasian wolves, suggesting that these ancient dogs likely came to the Americas with humans (Leonard et al., 2002). However, some ancient dogs in the Americas have mitochondrial haplotypes either shared with or nearly identical to those of North American wolves, suggesting either post-domestication admixture between dogs and wolves or even a separate domestication of canids in the Americas (Koop et al., 2000; van Asch et al., 2013). Ethnohistorical records indicate that Native American peoples used dogs as hunters, herders, haulers, sources of food, and companions, and some of these practices likely began

in prehistory (Schwartz, 1997).

Dogs have evolved to live with humans, and some of their adaptations provide a historical record of human activity as well. For example, recent studies have examined how genes governing starch digestion differ between dogs and wolves. Because dogs were domesticated before the advent of agriculture, it would have been important for them to adapt to the changing human diet and be more efficient at digesting starchy crops. One gene in particular, which codes for the enzyme alpha-2B-amylase (AMY2B), has two copies in wolves but as many as 30 copies in dogs (Axelsson et al., 2013). These differences in copy number correlate with the history of agriculture in a dog breed's region of origin (Freedman et al., 2014). Likewise, human populations have higher copy numbers of salivary amylase (AMY1) in regions with high-starch diets (Perry et al., 2007). Another example can be found in the Tibetan Mastiff, which has adapted to live with humans at high altitudes. A recent study has identified multiple candidate genes for this adaptation (Li et al., 2014), some of which (such as EPAS1, a transcription factor that regulates cell response to hypoxia) are the same genes that have been implicated in human high-altitude adaptation (Xu et al., 2011), and others (such as PLXNA4, a gene that promotes angiogenesis) that share similar functions (Scheinfeldt et al., 2012). Changes in genes expressed in the brain are also commonly found when comparing dogs and wolves (Saetre et al., 2004; Li et al., 2013), suggesting that behavioral differences between wolves and dogs have a genetic basis. These changes in behavior are thought to arise early in domestication (Kukekova et al., 2012).

Given the close bond that dogs and humans have shared throughout history, dogs can provide complementary data sources in studies of human populations. Notably, in cases where ancient human remains are inaccessible for use in genetic analysis, dogs can be used as a proxy to examine the population history of humans (Barta, 2006). Organisms that are closely involved with humans have likely moved across the earth following similar routes and at similar times, thus the genetic structure of these populations may reflect upon that of the humans they followed. Like dogs, rats have been distributed

worldwide by humans, and they have been used to trace worldwide migration patterns. For example, Polynesian rat populations have been used to inform multiple hypotheses about the origins of humans in Oceania (Matisoo-Smith and Robins, 2004).

Prior to AD 1492, dogs were widespread across the Americas and have likely been present since humans arrived on the continent (Leonard et al., 2002), so they potentially present an ideal study system for supplementing the reconstruction of human population history in the Americas. Dog remains have been recovered from the Jaguar Cave in Idaho and Agate Basin in Wyoming that date to over 10,000 ybp (Yohe and Pavesic, 2000), so DNA sequencing can be used to study the demographic history of dogs in the Americas spanning nearly 10,000 years. By studying dogs in parallel with humans, we may also learn more about the history of the peopling of the Western hemisphere than we can by studying humans alone, as dogs can help further test hypotheses about human migration. For example, because dogs have a shorter generation time (4 years) than humans (20-30 years) (Fuller, 1995), they have a higher substitution rate than humans, and theoretically they should have accumulated more variation, and possibly genetic structure, over a shorter timescale than possible with human genetic data (Walberg and Clayton, 1981). However, artificial selection (such as breeding) may result in reduced genetic variation or magnified population structure between regional populations.

Using genetic data from contemporary dogs alone to answer questions about the history of pre-contact Native American dogs can be problematic. Genetic studies indicate that much of the diversity of dogs in the early Americas has been lost after European contact (Castroviejo-Fisher et al., 2011). A broad sampling of modern village dogs in the Americas demonstrated that nearly all haplotypes identified were shared with dogs introduced to the Americas through European colonization. However, some isolated populations like the Carolina Dog and Alaskan Eskimo Dog show continuity with ancient samples and retain much of their indigenous diversity, with a maximum of 30% European admixture in the maternal line of extant populations (van Asch et al., 2013). Due to the collapse of Native American dog populations

following AD 1492 and the subsequent replacement of indigenous dog lineages during European colonization, sampling ancient dog remains may provide a much clearer picture of the ancient population history of dogs and their Native American domesticators in the early Americas.

Humans first entered the Americas roughly 15,000-20,000 years ago (Kemp and Schurr, 2010; Meltzer, 2010). Subsequent movements or expansions into North America from Northeast Asia followed the initial peopling before the Bering land bridge became submerged and separated Siberia from the Americas by 10,000-11,000 years ago (Forster et al., 1996; Tamm et al., 2007; Fagundes et al., 2008; Kemp and Schurr, 2010). The initial founders may have followed the Pacific Coast and rapidly spread southward, establishing the Monte Verde site in Chile, for example, approximately 15,000 years ago (Dillehay and Collins, 1988; Schurr and Sherry, 2004; Erlandson, 2007; Tamm et al., 2007; Fagundes et al., 2008). Expansion further inland likely occurred once glaciation withdrew and an ice-free corridor was opened (Hoffecker et al., 1993).

A topic of particular importance in studies of Native American population history is identifying founding mitochondrial haplotypes that were carried by the initial population that peopled the Americas. Accurate estimation of the initial diversity found in American populations is key to finding their geographic origin outside of the continents and estimating the founding population size. Torroni et al. (1993a) suggested the use of three criteria in determining which haplotypes of a haplogroup represent founding lineages. First, a founding haplotype is expected to be geographically widespread, cross-cutting linguistic and cultural divisions between Native American populations. Second, the founding haplotype should be central to the phylogeny of the haplogroup, as all other haplotypes in the haplogroup evolved from the founding haplotype, and are thus derived. Lastly, this founding haplotype should also be found in Siberia or elsewhere in Asia. Initially using these parameters, one founding haplotype was identified in each of the four haplogroups recognized at the time: A, B, C and D (Torroni et al., 1993b). Later, haplogroup X and D4h3a were identified as additional founder lineages using these criteria (Smith et al., 1999; Kemp et al.,

2007). Both of these haplogroups have been identified in ancient skeletal remains (Malhi and Smith, 2002; Kemp et al., 2007; Cui et al., 2013; Rasmussen et al., 2014), with D4h3a dating to at least 12,600 ybp. As argued by Kemp and colleagues (2007), mtDNA types observed in individuals of great antiquity in the Americas are likely to be founding lineages. A unique form of haplogroup M, observed in two ~5000 year old skeletons from the interior of British Columbia (Malhi et al., 2007), while not known to be widespread, may also represent a founder lineage.

More recently, mitogenome data has been used to infer that there were at least 14 founding mitochondrial lineages carried to the Americas: A2, B2, C1b, C1c, C1d, C4c, D1b, D1c, D1d, D2a, D3, D4h3, X2a, and X2g (Tamm et al., 2007; Achilli et al., 2008, 2013; Fagundes et al., 2008; Perego et al., 2009; Malhi et al., 2010; Hooshiar Kashani et al., 2012). Most of these haplogroups are geographically widespread. Notably, each of these haplogroups can be traced to a single ancestral haplotype that is derived by multiple substitutions relative to haplotypes present in Asia, suggesting a period of isolation for the Asia-to-Americas migrants from their source population (Tamm et al., 2007; Fagundes et al., 2008). This idea is commonly referred to as the Beringian Incubation Model (BIM) or Beringian Standstill Hypothesis, for which support can also be found in the nuclear genome (Tamm et al., 2007; Schroeder et al., 2009; Villanea et al., 2013).

Estimating the population size of the first humans to enter the Americas is also of particular interest. Changes in the effective population size of Native Americans have been estimated from the initial peopling of the Americas to the present using a Bayesian Skyline Plot (BSP) analysis incorporating a large number of complete mitogenomes from geographically and linguistically diverse populations (Kitchen et al., 2008; Mulligan et al., 2008). This analysis suggested a three-stage model for colonization, in which there was an initial period of divergence of the migrant population from their Central Asian source population, a period of 7,000-15,000 years of population stability in Beringia and a final period of rapid population expansion upon entry into the Americas, with a founder population with an effective size (N_e)

of 500-1,000 females. Critically, this model complements the BIM (Tamm et al., 2007) and suggests that while the initial founding population contained many founder haplogroups that diverged from Asian progenitor haplotypes during the Beringian occupation, the N_e of the founding population of the peopling of the Americas is surprisingly small. Given that the population history of dogs should be similar to that of humans, dogs may show a similarly small effective founding population size.

In this study, we have expanded the sampling of dogs in the early Americas by sequencing individuals recovered from three archaeological sites in North America. Using methods similar to those used for humans, we aim to characterize the population history of dogs in the early Americas by defining founding lineages and examining changes in the dog effective population size over time. In an attempt to move towards such a goal, in this study we sequence a portion of the hypervariable region (HVR) of mtDNA from ancient dog remains from these three archaeological sites. We then combine the sequence data with previously published sequence data from both ancient and modern dogs to identify founding dog mtDNA haplotypes in the Americas as well as infer population history of dogs in the Americas.

Methods

Sampling Information and Context

Samples were taken from three distinct archaeological sites across multiple temporal horizons. A map of the approximate location of archaeological sites from which dogs were previously studied, as well as the location of archaeological sites incorporated in this study can be found in Figure 2.1.

Janey B. Goode (JBG)

The Janey B. Goode site (11S1232) is a large, prehistoric settlement covering over six hectares in the American Bottom near Brooklyn, Illinois (Galloy, 2010). The site was occupied approximately 660-1350 ybp, and was most intensively occupied during the Terminal Late Woodland and Mississippian periods. Over 5,400 dog remains were recovered, with 103 individual dogs identified (Kuehn, pers. comm.). Approximately 80 dog burials were identified, with animals interred individually or in groups of

two or three near houses (Borgic and Galloy, 2004). Based on skeletal pathologies, most dogs were used as transport or pack animals. A subset of 39 individuals was identified for genetic analysis, and 35 of those samples were successfully extracted, amplified and sequenced to obtain mtDNA haplotypes.

Dionisio Point (DP)

Dionisio Point is a village site in coastal southwestern British Columbia, and includes two settlements: a large 5-plankhouse village (DgRv-3) that was occupied between 1500 and 1300 ybp and a single plankhouse (DgRv-6) that dates to between 1000 and 700 ybp. Substantial shell middens surround the houses at both sites. Extensive excavations have been completed at both sites over the last two decades (Grier, 2006; Grier et al., 2013) and abundant dog remains have been recovered from both house contexts and midden areas.

The sample of eight dogs we analyzed is derived from the shell midden behind the single, later plankhouse at DgRv-6. Dog remains are particularly abundant in this location. Articulated dogs were recovered from various midden layers, both in direct association with human burials and on their own. Isolated or fragmented dog remains were also frequently encountered in the deposits. Dogs of all ages are represented. The association of dog and human burials suggests more than haphazard deposition, and the midden may have been a highly symbolic place. An accurate minimum number of individuals (MNI) for the dogs represented at Dionisio Point has not yet been generated; a full analysis of the dog remains is in progress.

Albert Porter Pueblo (APP)

A large prehispanic Pueblo village in the central Mesa Verde region of southwestern Colorado, Site 5MT123 was occupied intermittently as early as 1400 ybp, but most of its occupation dates from the Pueblo II (1100-850 ybp) and Pueblo III (720-850 ybp) periods (Ryan, 2004). The two canid specimens sampled here were excavated between 2001 and 2004 by Crow Canyon Archaeological Center and are part of a larger on-going study of dog mtDNA variation in the American Southwest. Architectural details

and associated ceramic materials place these specimens within the 940-740 ybp interval (S. Ryan, personal communication).

DNA Extraction and Sequencing – University of Illinois Urbana-Champaign (UIUC)

Samples from the Janey B. Goode (JBG) site were extracted in a clean room environment dedicated to the extraction of DNA from ancient organisms at the University of Illinois Urbana-Champaign (UIUC), using a protocol developed previously by the Malhi lab (Cui et al., 2013). No modern dog samples have been processed in the ancient DNA lab, to ensure there is no cross-contamination between modern DNA and the ancient samples. Briefly, all teeth were wiped down with 6% sodium hypochlorite using a Kimwipe for at least a minute to remove surface contaminants, rinsed with molecular grade DNA-free water and dried under UV light and were drilled using a Dremel drill to produce 0.2 g powder. The powder was then digested in a solution of 4 mL EDTA, 300 µL 10% w/v N-lauryl sarcosine, and 100 µL 3.3% w/v proteinase K. The digestion was concentrated down using a centrifuge to a volume of 250 mL, and then extracted using the QIAQuick PCR Purification kit by Qiagen. All individual samples were extracted at least twice at different times to confirm all DNA sequences.

Multiple primers were used to amplify a portion of the hypervariable region of mtDNA (15421-15691 bp), as listed in Table 2.1 (Druzhkova et al., 2013). Samples were amplified using the polymerase chain reaction (PCR), with a mix as follows: 2 µL DNA, 13.25 µL molecular-grade water, 2 µL 10X PCR buffer, 1.2 µL 50 mM $MgCl_2$, 0.8 µL 100 mM dNTPs, 0.3 µL of each primer, concentration 20 mM, and 0.15 µL Platinum Taq DNA Polymerase (Life Technologies). The program used for PCR amplification involved an initial step at 94°C for two minutes, followed by 40 cycles of 94°C for 15 seconds, 55°C for 15 seconds, and 72°C for 12 seconds, with a final step at 72°C for 5 minutes, and successful amplification was verified with gel electrophoresis. Sanger sequencing of the PCR products was performed at the Roy J. Carter Biotechnology Center at UIUC. All individuals were sequenced at least twice for each extraction, and if a consensus was not reached between the extractions, a third extraction was performed to confirm the

DNA sequences of individuals. A consensus between two extractions, in which sequences amplified from multiple primers for each of two extractions matched exactly, confirmed the individual's sequence, and no individual required more than three extractions for sequence confirmation.

DNA Extraction and Sequencing – Washington State University (WSU)

Eight samples from DP and two samples from APP were extracted in the ancient DNA lab at WSU (Table 2.2), following the WSU method described in Cui et al. (2013). Samples were first tested for the presence of PCR inhibitors following Kemp et al. (2014) and subjected to repeat silica extraction until they were deemed inhibitor-free.

Two mitochondrial DNA fragments were PCR-amplified using the following primer sets: 1) D15401F (3'-AAGCTCTTGCTCCACCATCA-5') and D15595R (3'-GATATAATATTATGTACATGCTTAT-5'), 2) D15534F (3'-CTATGTACGTCGTGCATTAATG-5') and D15711R (3'-GGTTGATGGTTTCTCGAGGC-5'). Fifteen microliter reactions using Omni Klentag LA were conducted following Kemp et al. (2014) with an annealing temperature of 60°C. Successful amplification was confirmed via gel electrophoresis and amplicons were prepared and sequenced according to Kemp et al. (2014).

All sequences from this study are available on Genbank (accession numbers KJ189495-KJ189536).

Data Analysis

DNA sequences obtained from the JBG, DP, and APP dogs were combined with other ancient and modern North American dog and wolf mtDNA haplotypes reported in previous studies (Table 2.3). Ancient DNA haplotypes include samples from 8,000 year old dog burials in Siberia, which may have come from the same source population as ancient dogs in the Americas (Losey et al., 2013). Also used were ancient DNA haplotypes from Bolivia, Peru, Mexico, Argentina, Western Canada, and the United States (Koop et al., 2000; Leonard et al., 2002; Brown et al., 2013; Byrd et al., 2013; Thalmann et al., 2013). Modern haplotypes of North American and Eurasian wolves and all published dog haplotypes (as of 1/1/2014) were also used for comparison (Tsuda et al., 1997; Vila et al., 1997; Okumura et al., 1999; Savolainen et

al., 2002; Takahasi and Miyahara, 2002; Pang et al., 2009; Klütsch and de Caprona, 2010; Castroviejo-Fisher et al., 2011; Thalmann et al., 2013; van Asch et al., 2013). The DNA sequences obtained from the literature varied in length, so sequences were trimmed to produce a dataset that maximizes the number of individuals incorporated while showing variation along a shorter segment of 229 base pairs [nucleotide positions (nps) 15458-15687, according to Genbank Accession NC002008] of the mitochondrial genome.

Putative founding haplotypes were identified using the following criteria, modeled after similar criteria used for inferring founding haplotypes of Native Americans (Torroni et al., 1993a). A founding haplotype should be present in multiple geographic regions of the Americas and should be central to a phylogeny of dog mtDNA sequences. Additionally, a founding haplotype may be found both in the Americas and in Asia. However, if a haplotype is found to be infrequent or geographically localized in the Americas and also differs by multiple substitutions from other dog haplotypes, it may also be a putative founding haplotype. It would be more likely for a sequence that differs by five or more substitutions from other founding haplotypes to be another founding haplotype than for it to be derived from a much more distant haplotype.

The dataset of ancient dog and modern dog and wolf sequences was aligned using Bioedit, and the program Network was used to construct a median-joining network for the DNA sequences (Bandelt et al., 1999). A network is a visual representation of how haplotypes relate to one another and is useful to determine if certain haplotypes are shared among populations. The median-joining method constructs trees that minimize the genetic distance between haplotypes by linking clusters of closely-related sequences, and then resamples those clusters to produce the most parsimonious network (Bandelt et al., 1999). For each geographic region studied as well as for the wolf samples, additional calculations were made using Arlequin (Excoffier and Lischer, 2010) to compare diversity within and between groups. For each group, the number of haplotypes and θ_s, the number of segregating sites in the sample corrected by the number of individuals in a sample (Watterson, 1975) were calculated for each group. Nucleotide

diversity was calculated for each group as a measure of sequence diversity weighted both by haplotype frequency and the number of substitutions that differ from other haplotypes in the group. These measures of genetic diversity are often used to compare populations and estimate how populations differ from one another. A population with lower diversity might indicate a bottleneck event or deliberate breeding, and a population with higher diversity might indicate a larger and more stable population size over a longer period of time or a population that experienced gene flow. Since only two samples were analyzed from the Albert Porter Pueblo population, they were omitted from this analysis. In some cases, such as in Alaska and Mexico, samples are derived from multiple archaeological sites, in which case measures of diversity were also calculated for each archaeological site containing multiple dogs. Finally, an analysis of molecular variance (AMOVA) was also performed to estimate groupings that explained the most variation, to estimate what populations were most closely related.

Some mtDNA haplotypes were identified as "outliers". These "outliers" were ancient dog haplotypes that differed from other ancient dog haplotypes by at least four substitutions, and are either haplotypes shared with wolves or putative founding haplotypes due to their genetic distance from the other sequences. In addition to what was reported in the literature, these haplotypes were identified as wolf or dog haplotypes by constructing a distance tree. Using all modern dog haplotypes, all published wolf haplotypes (as of 1/1/2014) and all ancient dog haplotypes, a phylogeny was constructed in MEGA, using a maximum likelihood criterion with an HKY+I+G model of substitutions (Tamura et al., 2011). This method of tree construction starts with a neighbor-joining tree (incorporating sequences into the tree one at a time in order of similarity until all are part of the tree), with branch lengths and substitution rates optimized to their maximum likelihood values to produce the final tree.

A dataset consisting of ancient dog haplotype sequences not closely related to wolf haplotypes (i.e., those not likely of wolf-dog admixture) and all modern dog haplotype sequences found in dog breeds originating in the Americas was analyzed under a Bayesian coalescent framework. This analysis was

performed using the Bayesian Markov chain Monte Carlo (MCMC) methods implemented in BEAST (version 1.8.0, Drummond et al., 2012). For this analysis, an HKY+G+I substitution model, a strict molecular clock, and an extended Bayesian skyline plot demographic model (with linear changes in population size; Heled and Drummond, 2008) was used to infer the historical dynamics of dog populations. Importantly, the family of Bayesian skyline plot demographic models provides a means to infer past population histories using both ancient and contemporary genetic samples without a priori definition of a parametric model of past population dynamics (i.e., constant or exponentially growing population size). Mean dates for archaeological sites were used as sampling dates (Table 2.3 and Table 2.4) of the aDNA sequences and provided independent calibrations for the molecular clock; a CMTC prior was used as a prior for the substitution rate (Ferreira and Suchard, 2008). Unless otherwise noted, default priors and operator values were used. Markov chains were run for 200 million generations with samples taken every 10000 generations; convergence of three independent MCMC runs was assessed using Tracer (version 1.6, Rambaut and Drummond, 2007), and all MCMC samples were combined after the first 10% of samples were discarded as burn-in. The combined data files were used for final inferences about the historical population dynamics of Native American dogs and to produce a summary genealogy of the HVR1 sequences in this dataset.

To assess the accuracy of our coalescent inference and determine if our punctuated sampling of ancient dog HVR1 sequences introduced a bias to our analyses, two sets of simulations were performed. First, ten subsets of the dataset consisting a random sample of 25 aDNA sequences across all times were also analyzed in BEAST using the same models and priors from the analysis of the full dataset. This was done to assess any possible bias in the estimation of past population dynamics introduced by both non-random space (i.e., the inter-relatedness of dogs at sites) and time (i.e., multiple samples come from the same horizon). Second, we performed simulations to produce synthetic datasets with a known constant demographic history and substitution and clock model parameters identical to those estimated in the

analysis of the full dataset. These simulations were performed in BEAST (Bielejec et al., 2014) and analyzed with the HKY+G+I substitution, strict clock, and extended Bayesian skyline plot demographic (linear variant) models. One hundred simulations were performed to assess whether punctuated sampling of low information genetic data might produce artifacts in reconstructed skyline plots.

To compare shifts in dog population size with shifts in human population size, an extended Bayesian skyline plot was constructed using the dataset employed in O'Fallon and Fehren-Schmitz (2011), which consists of living individuals representing all five Native American haplogroups and ancient individuals from Ontario, Illinois, and Peru, and is thus similar to our dataset in having both modern and ancient sequences. To allow for direct comparison, only the HVR1 of these individuals was used in the analysis. This analysis was also performed in BEAST 1.8.0 using an HKY+G+I substitution model, an uncorrelated lognormally distributed relaxed molecular clock (Drummond et al., 2006) with a fixed mean substitution rate of 1.64×10^{-7} substitutions/site/year (Soares et al., 2009) and an extended Bayesian skyline plot demographic model. Sampling dates for the ancient individuals were taken from the original analysis (O'Fallon and Fehren-Schmitz, 2011). To assess the accuracy of demographic reconstruction using short HVR1 sequences, 100 simulations were performed in which synthetic datasets were produced from known demographic history and analyzed in BEAST. Synthetic data were simulated using the same substitution model parameters estimated in the empirical analysis and under a demographic history similar to that inferred from the empirical data: an initial population size of 10^3 at 15,000 ybp, and a sudden decrease in population size to 5×10^4 ybp. These synthetic datasets were analyzed with the HKY+G+I substation, strict clock (rate = 1.64×10^{-7} substitutions/site/year), and extended Bayesian skyline plot demographic (linear variant) models. All graphs were produced using R.

Results

The individuals sequenced in this study represent a total of nine different HVR1 haplotypes, four of which are novel. Novel haplotypes were authenticated by confirming identical sequences from two extractions, as well as by amplifying and sequencing each extract twice, to ensure that the novel

substitutions are not due to miscoding lesions, a common problem in ancient DNA analysis (Gilbert et al., 2003). All haplotypes are shown in Table 2.4 – due to the short length of the sequence, nearly all of the individuals with haplotypes that are not novel have identical sequences to multiple modern dog haplotypes. The measures of genetic diversity are shown in Table 2.5 for all geographical regions, and in Table 2.6 are subdivided further by archaeological site. Taking a regional perspective, Alaska and Mexico have the highest diversity, followed by California and Western Canada. Siberia, Bolivia, Illinois (represented by JBG) and coastal British Columbia (represented by DP) have much lower levels of diversity. When subdivided into archaeological sites, Tula, Mexico, has the highest diversity, followed by Western Alaska and California – the diversity at the Fairbanks, Alaska site is much lower.

The network in Figure 2.2A shows all published ancient dogs of the Americas compared to wolf haplotypes worldwide. The addition of samples from Janey B. Goode, Dionisio Point and Albert Porter Pueblo introduce six new mitochondrial haplotypes not found previously in dogs from the Americas. One haplotype is found to have a particularly high frequency in the sample, and is represented by twenty-eight dogs from Janey B. Goode, as well as one dog each from Illinois, Florida, California, Albert Porter Pueblo, and Siberia as well as one Asian wolf (a similar network color-coded by geographic region is presented in Figure 2.3). Notably, many of the other ancient dog haplotypes differ from this common haplotype by only one or two nucleotide substitutions. There is also some regional clustering of haplotypes from ancient dogs – nearly all haplotypes from South American dogs cluster closely together and differ only by a few substitutions, whilst most of the Janey B. Goode haplotypes only differ by a few substitutions as well. Interestingly, there are also a few clusters of wolves, highlighting the distinctiveness of the wolf versus dog haplotypes. However, some dog haplotypes are nearly identical to wolf haplotypes, and are separated by only one or two substitutions. Specifically, most haplotypes from the Canadian dog sample are shared with or only slightly distant from a North American wolf haplotype, whilst a Mexican dog haplotype is identical to a European wolf haplotype that differs by only a single substitution from a North American

wolf haplotype.

A second network was constructed that consists solely of ancient dog samples, to provide an illustration of how the dog haplotypes relate to one another without incorporating the wolf haplotypes, and is represented by Figure 2.2B. It is more apparent here that there are multiple "outlier" haplotypes. As discussed above, four of the Western Canadian dogs and the Mexican dog in the figure are likely the result of admixture between dogs and wolves. These wolf-like lineages may indicate that dogs were separately domesticated in the Americas. A single haplotype represented by an Alaskan dog seems dissimilar to both dogs and wolves on the network in Figure 2.2A, but all individuals with more than four substitutions different from other dog haplotypes in Figure 2.2B are considered to be "outliers".

A phylogeny was constructed to infer the relationship of the outlier sequences to all published dog and wolf haplotypes (Figure 2.4). Over all, the wolf and dog haplotypes seem to be mixed throughout the tree. The haplotypes from the ancient dogs that are identified as "outliers" in Figure 2.2B are likely (i) founding haplotypes present in the dog population, (ii) haplotypes that show admixture with North American wolves, or possibly (iii) the result of a separate domestication of dogs in the Americas. Notably, multiple outlier haplotypes are more closely-related to wolves than dogs, suggesting admixture or independent domestication. One haplotype from an Alaskan dog is not closely-related to modern dog or wolf haplotypes, whilst two other outlier haplotypes cluster more closely with modern dog haplotypes than wolf haplotypes. Interestingly, the latter two outliers cluster with some haplotypes that are only found in Siberia, the Americas or Eastern Asia (Pang et al., 2009), possibly indicative of a northeast Asian origin, consistent with the origin of Native Americans (Forster et al., 1996).

The extended Bayesian skyline plot of the dogs shows a relatively stable population from the present to the time of the most recent common ancestor, ~9000 ybp, with the exception of a small dip in population size around 1000 ybp (Figure 2.5A). A summary genealogy relating the dog samples in the dataset produced from the posterior distribution of genealogies sampled in the Bayesian coalescent

analysis is presented in Figure 2.6. The simulations using random subsets of the ancient samples shows no decrease in effective population size over time, suggesting that the population decline is likely an artifact of the sampling, as shown in the composite extended Bayesian skyline plots in Figure 2.6. Furthermore, the simulations in which synthetic datasets produced under a constant population size model and identical sampling distributions across time (i.e., sequences of the same age) also show a consistent signal of a stable population over time without an event near 1000 ybp (Figure 2.5B). We also produced extended Bayesian skyline plots for Native Americans that allow us to contrast the demographic history of the first Americans with that of their dogs. The extended Bayesian skyline plot of the human samples shows an increase 15,000 to 20,000 ybp with a slight decrease at ~500 ybp (Figure 2.5C). The simulations of synthetic Native American HVR1 datasets shows a similar population increase at ~15,000 ybp, but a much smaller or even absent decrease toward the most recent times (Figure 2.5D).

Discussion

This study of the historical dynamics of Native American dogs has significantly increased the size and geographic diversity of genetic data from ancient dogs in the Americas. Of the nine haplotypes identified in the populations from JBG, APP, and DP, four are novel, which suggests that additional variation present in dogs in the Americas prior to European contact has yet to be identified. Additionally, our ancient dog sample from the Janey B. Goode site is the largest sample of dog mtDNA data from a single archaeological site. The fact that so many dogs were buried at this site suggests that the people who lived there placed considerable value on their canine companions. Critically, the large sample size also increases the likelihood that estimates of genetic diversity more accurately reflect the true genetic diversity of this dog population.

Notably, with the increased sample size and expanded geographic distribution enabled us to identify putative founder haplotypes. Leonard et al. (2002) identified five putative founding haplotypes, two of which were present in ancient dogs and three of which were found in contemporary dogs, as well

as a clade “a” consisting of haplotypes of Latin American dogs unique to the Americas. These haplotypes were identified using a different nomenclature, such that some of these haplotypes have an identical hypervariable region to multiple dog haplotypes identified in Pang et al (2009), which can make the founding haplotypes difficult to identify in some cases. According to guidelines established by Torroni et al. (1993a), haplotypes that fit these guidelines and are likely putative founding haplotypes are identified with an asterisk in Figure 2.2B. Given that the most common haplotype is shared between dogs from the Midwest, Southeast and Southwest, as well as Siberia (the probable source of the American dog populations; Forster et al., 1996; Leonard et al., 2002), this haplotype is likely a founding haplotype. Another haplotype, one shared with dogs from JBG, Alaska, and Peru, and identified by Leonard et al. (2002) as a putative founding lineage, is likely also a founder haplotype because it is widespread across the Americas. A third haplotype, also identified by Leonard et al (2002), is shared between three Alaskan dogs and one dog from JBG, and could also be a putative founding lineage. Given that these founding haplotypes are characterized by only a portion of the mtDNA HVR, the haplotypes cannot be identified definitively unless complete mitogenomes of these dogs are sequenced. The clade “a” identified by Leonard et al. (2002), however, seems to be a subhaplogroup localized to South America.

There are multiple outliers in the network that are distantly related to most of the other dog haplotypes from the early Americas. Some of them are closely related or identical to wolf haplotypes, as shown in Figures 2.2A and 2.2B. These samples could indicate admixture with North American wolves or a separate domestication (or events) from North American wolves. This is most clearly demonstrated by the Western Canadian haplotypes, which are almost identical to a North American wolf haplotype. Interestingly, the Mexican sample has an identical haplotype to a European wolf, but that haplotype is also only one base pair different from a North American wolf haplotype. Given that mtDNA only captures a fraction of an individual’s ancestry (i.e., maternal only) and only a portion of the hypervariable region of the mitochondrial genome was analyzed in our study, our results do not provide the resolution necessary

to distinguish between the two possibilities of admixture and separate domestication in the Americas. Sequencing of other regions of the genome would be required to determine if these sequences derive from a single female wolf that interbred with domesticated dogs in the Americas, or if this haplotype represents the ancestor of domestic dogs in the Americas.

There is a total of eight outlier haplotypes shown in Figure 2.2B that differ significantly from other ancient dog haplotypes, and there are multiple reasons why such outliers might be present in our dataset. First, these haplotypes represent dogs that came to the Americas via Siberia that exhibited haplotypes different from the other founding haplotypes, and represent additional founder lineages. Second, these dog haplotypes are shared with wolf haplotypes that have either gone extinct or have yet to be sampled and published in the literature. If they are shared with wolves, admixture or separate domestication are both viable possibilities. Interestingly, some of these haplotypes cluster more closely with dog haplotypes, suggesting that these dogs represent different founding haplotypes – they are likely too distantly related to the other haplotypes in the Americas to have diverged from another founding haplotype, but are more closely-related to contemporary dog haplotypes than wolf haplotypes. Notably, some of the modern dog mtDNA haplotypes that are most closely related to the "outliers" in Figure 2.3, such as A31, A121, and B28, are exclusively found in Asian dogs, further suggesting that these were founding haplotypes (Pang et al., 2009). However, it is surprising that wolf and dog haplotypes are so thoroughly intermixed in Figure 2.3, when in Figure 2.2 the wolf haplotypes seem far more distant from the dog haplotypes. The cases of admixture identified in the Network explain some of this intermixing, but the similarities may also be due to the use of only the hypervariable region in this study. The relatedness between the dog and wolf haplotypes may be much lower than the tree indicates if the complete mitogenome were incorporated into the tree.

When comparing the two measures of genetic diversity (theta S and nucleotide diversity), the values are strongly positively correlated (R^2=.66, p=.0024, data not shown) across populations, though the

estimates for each statistic differ by two orders of magnitude. Although these statistics are expected to estimate roughly the same value for a population at equilibrium, they were calculated on a per sequence (theta S) and a per nucleotide position (nucleotide diversity) scale; when correcting nucleotide diversity for the number of nucleotide positions in the HVR1 fragments used, the estimates converge on the same scale. Interestingly, the correlation is stronger when excluding the sample from Tula, Mexico, which has a high theta S estimate relative to its estimated of nucleotide diversity. The dogs from Tula, Mexico, have mitochondrial sequences that differ greatly from one another in a very small sample size, which accounts for this difference and is unusual amongst the regional datasets.

Surprisingly, the genetic diversity measurements of the dog populations studied do not seem to follow any geographic pattern. For example, the Alaska, Western Canada, Mexico, and Peru populations display high levels of diversity, whereas the sites in Bolivia, and the North American Southwest and Midwest have much lower levels of diversity. This may be due to sampling – some of the regions used for analysis had multiple archaeological sites, which might inflate the genetic diversity as compared to regions with one or few. However, as shown in Table 2.5, even when samples from only individual archaeological sites are used in the analysis, diversity levels remain high in Mexico and Alaska, while Illinois and coastal British Columbia have low levels of diversity. The measures of nucleotide diversity for JBG and DP are also lowered, reflecting the high frequencies of common haplotypes found at these sites, though diversity estimates remain low even when the haplotype frequency bias is removed because all of the haplotypes sampled from each site are closely related. This could indicate that the population arose from a small number of closely-related female founders, or that variable breeding regimens of domesticated dogs were practiced very early across the Americas.

Examining the skeletal remains for morphological similarities or similar skeletal modifications could further support the possibility of selective breeding. Measurements of femori and humeri can be used as a reliable proxy for carnivore size (von Valkenburgh, 1990). Femoral and humeral length in the

JBG individuals have standard deviations of only 8 and 6 mm, respectively, indicating phenotypic homogeneity. This homogeneity could be explained by inbreeding or selection, whether natural or artificial. Observations of village dogs that scavenge at human settlements but don't directly interact with humans suggest that they are all very similar in size, likely because a balance must be struck between being small enough to thrive on limited nutrients and being large enough to defend against other dogs (Coppinger and Coppinger, 2001). Humans could also be directly selecting for a particular size of dog as well; given that JBG dogs were likely used for hauling supplies (Borgic and Galloy, 2004), perhaps the humans living there were selecting for dogs well-suited for hauling a given weight for long distances. An inbred population would be highly genetically similar and therefore phenotypically similar as well. Additionally, the dogs sampled from the DP site in this study come from a single shell midden, which represents only a small portion of the dogs recovered from the DP site that are available for study (Barta, 2006). These dogs could have been closely related and were therefore buried in the same shell midden, but the dog population at DP could have been much more genetically diverse overall. In future studies, comparing the phenotypic measurements and genetic data from other dog populations to genetic data will help determine the correlation between phenotypic and genetic variability in Native American dogs.

We performed Bayesian coalescent analysis to estimate the historical demography of dogs in the Americas using the methods implemented in BEAST. The extended Bayesian skyline plot (EBSP) shows a stable dog population size across time, with a small dip around 1000 ybp. However, our subsequent analyses suggest that the genetic signal of the EBSP is biased by the sampling. Specifically, large numbers of dogs were sampled from the same time period with identical haplotypes (mostly at JBG), and this likely causes the "dip" in the plot roughly 1,000 ybp. We have two lines of evidence that support this conclusion. First, we analyzed simulated datasets consisting of random subsets of ancient samples (all modern samples were retained), which theoretically eliminated large numbers of duplicate sequences from the same ancient sampling period, which produced EBSPs with no discernible decrease in population size

(Supplemental Figure 2.1B). Second, our analysis of synthetic datasets simulated under a constant population size and with the same sampling regime (i.e., with the same sampling dates) but without any correlation between sampling time and relatedness (i.e., samples from the same period were not more closely related than those from other sampling periods) also produced EBSPs without any consistent deviation from a constant Population (Figure 2.4B). Combined, these simulations suggest that the decrease in population size at ~1000 ybp is an artifact introduced by sampling bias (i.e., the correlation between sampling time and relatedness).

As posited above, there are many reasons why we should expect human and dog population dynamics to be correlated, as the histories of dogs and humans are intertwined. To investigate human history of the Americas at the same genetic resolution used in our analysis of dog history, we performed Bayesian coalescent analysis of the human HVR1 data used in O'Fallon and Fehren-Schmitz (2011) using the same demographic model we applied to the analysis of the dog HVR1 dataset. As expected, the human EBSP shows an increase ~15,000 ybp, followed by a stable population size lasting nearly to the present day, with a late, non-significant dip in effective population size near the present. Interestingly, this EBSP differs from the EBSP produced from the complete human mitogenome, in which there is a clear population decline around the time of European contact (O'Fallon and Fehren-Schmitz, 2011). Our analysis of synthetic data simulated under a demographic model with a recent-post-Columbian population collapse of 50% also did not produce EBSPs that consistently reflect this event. Combined, these contrasting results suggest that the hypervariable region of mtDNA may not be expected to reveal fine-grained changes in recent population history, though it may be possible that larger samples of HVR1 sequences might contain enough signal to reliably reflect recent population histories.

When considering both the human and dog EBSPs, the population stability found in the dog plot is unexpected, as one would expect the dog population to increase in size over time, as the human population did. Intriguingly, this unexpected finding might indicate that humans were controlling dog

matings and effectively breeding them, or that the population of dogs in the Americas quickly reached long-term carrying capacity. The stable population of dogs had a median effective population size of ~1,000 female individuals at the time of the first dog burial in the Americas. This is consistent with other estimates of dog effective population size that suggest a global population of roughly 10,000 female dogs at the same time (Thalmann et al., 2013). Additionally, the coalescence dates estimate for our dataset of American dogs is also surprising, given that by ~9000 ybp dogs should have been established across the Americas as they are thought to have arrived in the Americas ~15,000 ybp with humans. However, it is interesting that from 9000 ybp to the present, the population of dogs in the Americas roughly mirrors that of humans, in that both are stable for long periods of time.

Importantly, sequencing the hypervariable region alone captures a lot of diversity in a short stretch of sequence, but does not provide a full picture of mitogenome diversity. As demonstrated by Table 2.4, multiple dog haplotypes have identical sequences in the region we studied, so it is possible that we are underestimating the diversity of dog mitogenomes present in the Americas, as was found true of human mitochondrial diversity as mitogenome sequencing became more routine (Tamm et al., 2007; Achilli et al., 2008; Fagundes et al., 2008). This underestimation, combined with the possibility of breeding practices mentioned above, may mean that some of the putative founding haplotypes we have identified are not as frequent or as widespread as our results seem to indicate. Furthermore, if such population structure does exist, widespread sampling of dogs from multiple locations will be necessary to accurately characterize the diversity of dog founding mtDNA lineages. Calculating measures of population diversity using a short segment of DNA has revealed some interesting differences in diversity between populations, but it is pertinent now to sequence complete mitogenomes of these dogs to ensure that short nucleotide sequences have not biased these estimates. Ultimately, these analyses have confirmed that using complete mitogenomes can also provide a clearer picture of dog population history.

Given that dogs and humans have lived interdependently for thousands of years, dogs have

potential for use as a proxy to test human models of migration. Comparison of human populations with the dog populations in this study could demonstrate similarities that show dogs' utility as a complementary dataset for testing migration models. Notably, the JBG site also has human remains, and DNA analysis of these individuals is currently in progress. However, the human remains postdate the dog remains by at least 200 years at this site, and so the populations may not be directly comparable. In the cases of APP and DP as well, there are no human remains that have been found to be contemporaneous with the dog remains sampled in this study. However, this absence of human remains highlights the utility of using dogs as a proxy to learn more about how people lived in regions and time periods from which no human burials have been found.

Additionally, the haplotype distribution and diversity of dog haplotypes in a given region can allow us to infer human interactions with dogs, which can tell us more about the human population. Deliberate burial of dogs indicates that humans took care of the animals, and low levels of diversity can suggest deliberate breeding practices. If the dogs were being bred for a specific purpose, such as to haul sledges as in the case of the JBG dogs, it is possible that their owners were selecting for specific traits. If a single haplotype is frequent in multiple regions, it is possible that dogs migrated with humans to multiple locations. Haplotype distributions can also reveal instances of genetic drift, as seems to be the case in South America. Of the nine individuals currently sequenced from South America, eight of them are all part of the same derived clade. This founder effect has also been identified in human populations (Wang et al., 2007), supporting the idea that dogs and humans have traveled together and their populations have changed over time in similar ways.

Conclusion

Combining our data generated in this study with all published ancient dog haplotypes has revealed new diversity and multiple shared haplotypes across broad regions of the Americas. Some archaeological sites exhibit low levels of genetic diversity, suggesting the possibility of deliberate breeding practices in

the area. The most frequently identified haplotype in the sample is likely a founding haplotype, especially since it is identical to that of an ancient Siberian dog, possibly from the same source population. Two additional haplotypes that are not as frequent or as widespread, but still had a broad sampling, are also putative founding haplotypes. Additional "outlier" haplotypes that are infrequent but share similarities with modern dog haplotypes from Eurasia may also be possible putative founding haplotypes. This provides a minimum of three to five founding haplotypes. Given that each of these haplotypes are identical in sequence to multiple contemporary dog haplotypes, determining the nomenclature of the founding haplotypes should likely incorporate whole mitogenome sequences. Of the ~273 dog haplotypes that have currently been identified, 29 can be found in dogs of pre-European contact Americas.

Multiple DNA sequences were identified that were identical to haplotypes of North American wolves, and could represent admixture with wolves or a separate domestication event(s). However, the hypervariable region of the mitochondrial genome lacks power to draw certain conclusions (i.e., poor resolution of fine haplotype structure) and as mitochondrial DNA represents only the direct maternal line of numerous ancestors, it is not possible to distinguish between admixture with wolves and a separate domestication event. Examination of other regions of the genome, such as autosomal SNPs or even the exome (the complete coding region of the genome) could better elucidate the evolutionary history of these dogs, as they are a result of multiple ancestral lineages.

The analysis of mitochondrial DNA haplotypes in this study provides insight to the population history of dogs in the Americas and brings us closer to a comparison between population histories of Native Americans and their dogs. However, the hypervariable region analyzed in this study is only a short segment of the mitochondrial genome and likely subject to recurrent mutations which can obscure identifications of founding haplotypes in the Americas (Malhi et al., 2002). Additionally, the high frequency of specific dog haplotypes in certain archaeological sites, if the result of breeding practices, can bias estimates of diversity as shown in the extended Bayesian skyline plot estimates of dog population

histories. Sequencing complete mitochondrial genomes in pre-European contact dogs will likely provide a less biased view of mtDNA diversity of dogs in the early Americas. For example, although the dog dated to 9000 ybp from Illinois and the dog uncovered from Florida dated to 1000 ybp share a hypervariable region haplotype, their complete mitochondrial genomes differ at twelve nucleotide sites (Thalmann et al., 2013). Given that their shared haplotype is the most numerous in ancient American dogs sampled thus far, it will be important to sequence additional complete mitochondrial genomes for comparison purposes and to gain a more nuanced view of the shared history of dogs and humans in the Americas.

Figures and Tables

Figures

Figure 2.1: Map depicting the locations of all archaeological sites in North and South America containing dog remains from which mtDNA sequences were obtained. All black dots indicate sites from the literature, and blue dots indicate sites added in this publication. Numbers next to each dot indicate population size, and multiple numbers next to a single dot indicate that multiple archaeological sites in a single location were sampled.

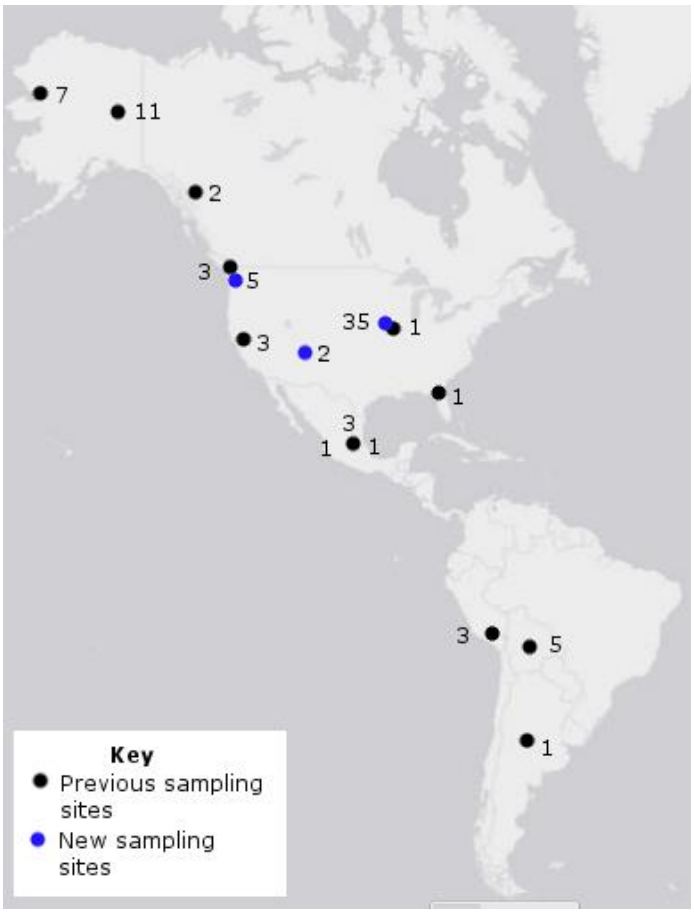


Figure 2.2: Mitochondrial haplotype networks of ancient dogs in the Americas. "Arctic dogs" are considered to be individuals from archaeological sites in Western Canada and Alaska. Each color node represents a haplotype, a unique mitochondrial DNA sequence. The size indicates the number of individuals with the particular haplotype. The length of the lines connecting the nodes indicate the number of substitutions that separate any two haplotypes. Small red diamonds indicate uncertainty – in these cases it is impossible to tell the exact order of the substitutions in question. **A)** Network of ancient American dog haplotypes with ancient and modern wolf haplotypes, including those published in this study. **B)** Network of ancient American dog haplotypes, including those published in this study. Asterisks by a haplotype indicate that it a putative founding haplotype.

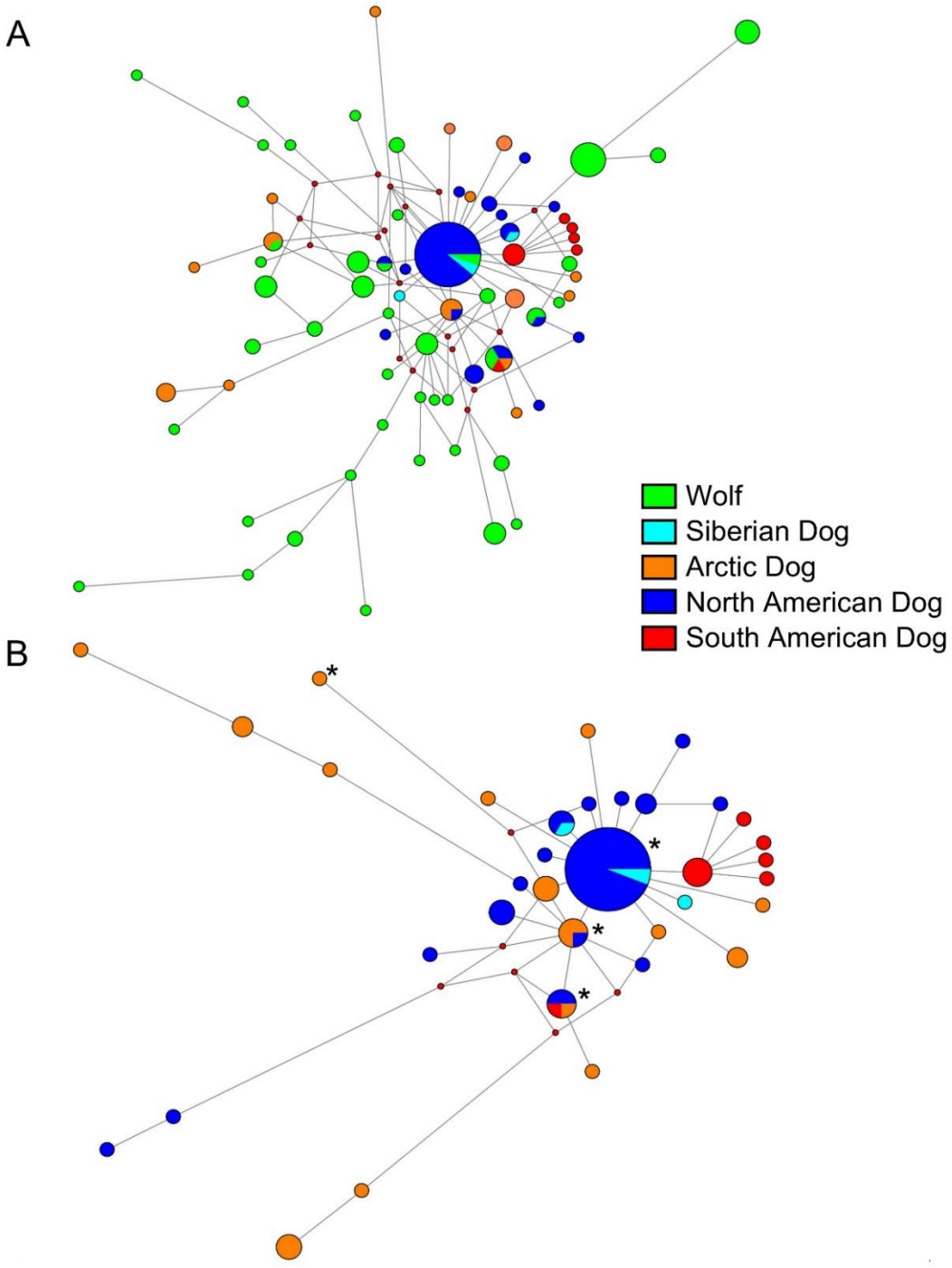


Figure 2.3: A network of all published ancient dog and ancient and modern wolf haplotypes, including the ones added in this study, with individuals color-coded by region. The size indicates the number of individuals with the particular haplotype. The length of the lines connecting the nodes indicate the number of substitutions that are different between the two haplotypes. Light grey circles indicate discrepancies– it is impossible to tell the exact order of the substitutions in question, so there are some ambiguities.

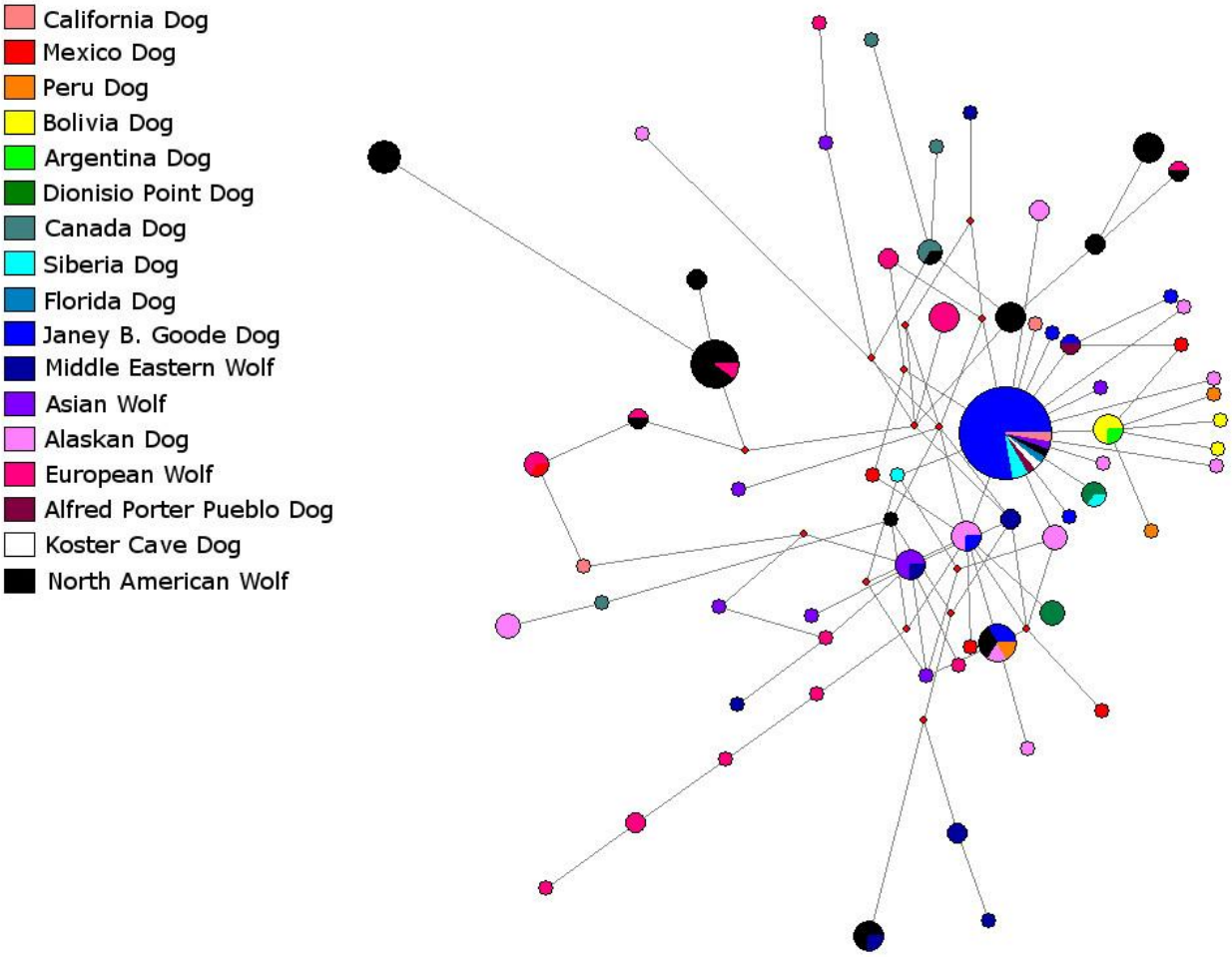


Figure 2.4: Phylogeny of modern and ancient dogs and wolves of the Americas. If multiple ancient dogs, modern dog haplotypes or wolves from a particular region shared a branch on the tree, they were omitted to simplify the tree for presentation. Triangles represent wolf haplotypes, squares represent modern dog haplotypes, circles represent ancient dog haplotypes and diamonds indicate "outlier" ancient dog haplotypes that differ by at least five substitutions from other founding haplotypes identified in this study, as shown in the key. All samples that are neither contemporary dog haplotypes nor part of this study are marked with a Genbank accession number belonging to the individual sampled or one with an identical sequence.

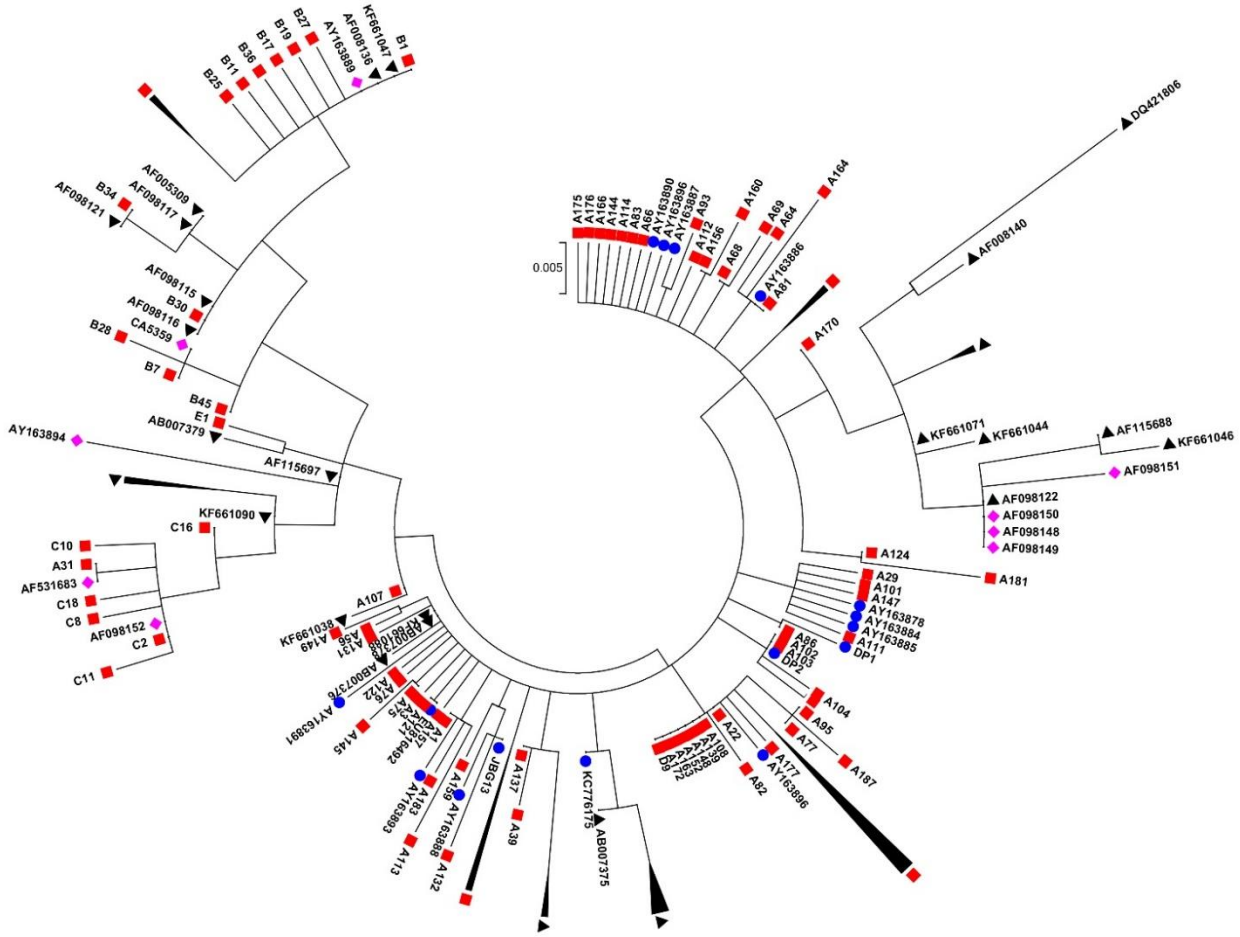


Figure 2.5: Extended Bayesian skyline plots of the dog mtDNA hypervariable region compared to those of the human hypervariable region, using logarithmic values for effective population size standardized by generation time as the y-axis and time as the x-axis. **A)** EBSP for the dog mtDNA HVR, using a strict molecular clock and an HKY+I+G model. **B)** Plot of median estimates from EBSP analyses of 100 simulations of dog-like HVR1 data in a constant-size population, using the same substitution and clock models and parameters used in the analysis of the empirical data, including sampling dates, but with no correlation between sequence sampling time and relatedness. **C)** EBSP for the human mtDNA HVR, using a relaxed molecular clock and an HKY+I+G model. **D)** Plot of median estimates from EBSP analyses of 100 simulations of Native American-like HVR1 data in a population with a population increase 15000 ybp and a crash 500 ybp, using the same substitution model and parameters estimated in the analysis of the empirical data, including sampling dates, and a strict molecular clock.

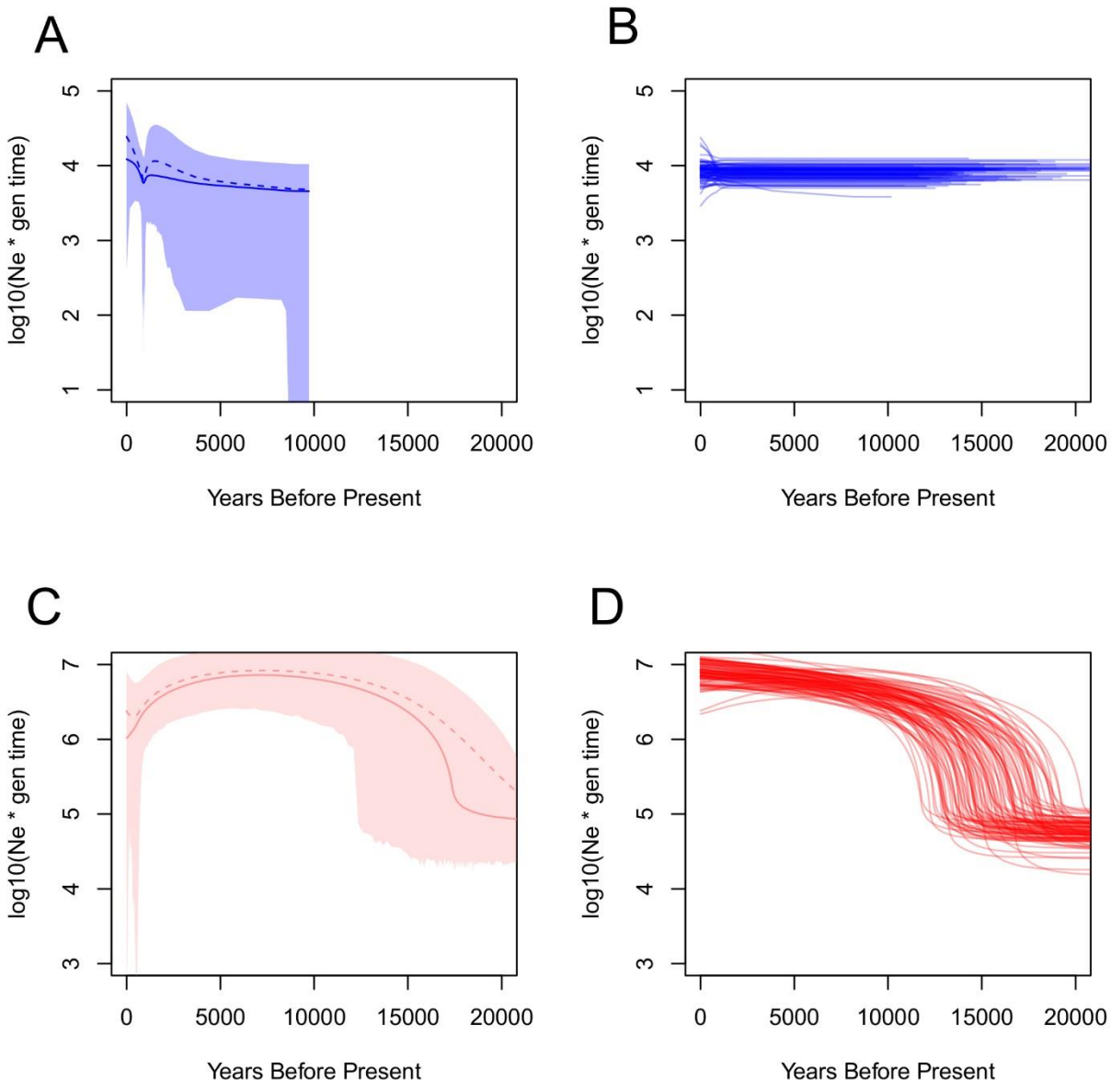


Figure 2.6: Extended Bayesian skyline plot for the dog mtDNA hypervariable region using logarithmic values for effective population size standardized by generation time as the y-axis and time as the x-axis. **A)** EBSP for the dog mtDNA HVR, using a strict molecular clock and an HKY+I+G model. **B)** A composite of Bayesian Skyline Plots for ten simulations of the dataset, incorporating a random sample of 25 ancient individuals in each simulation.

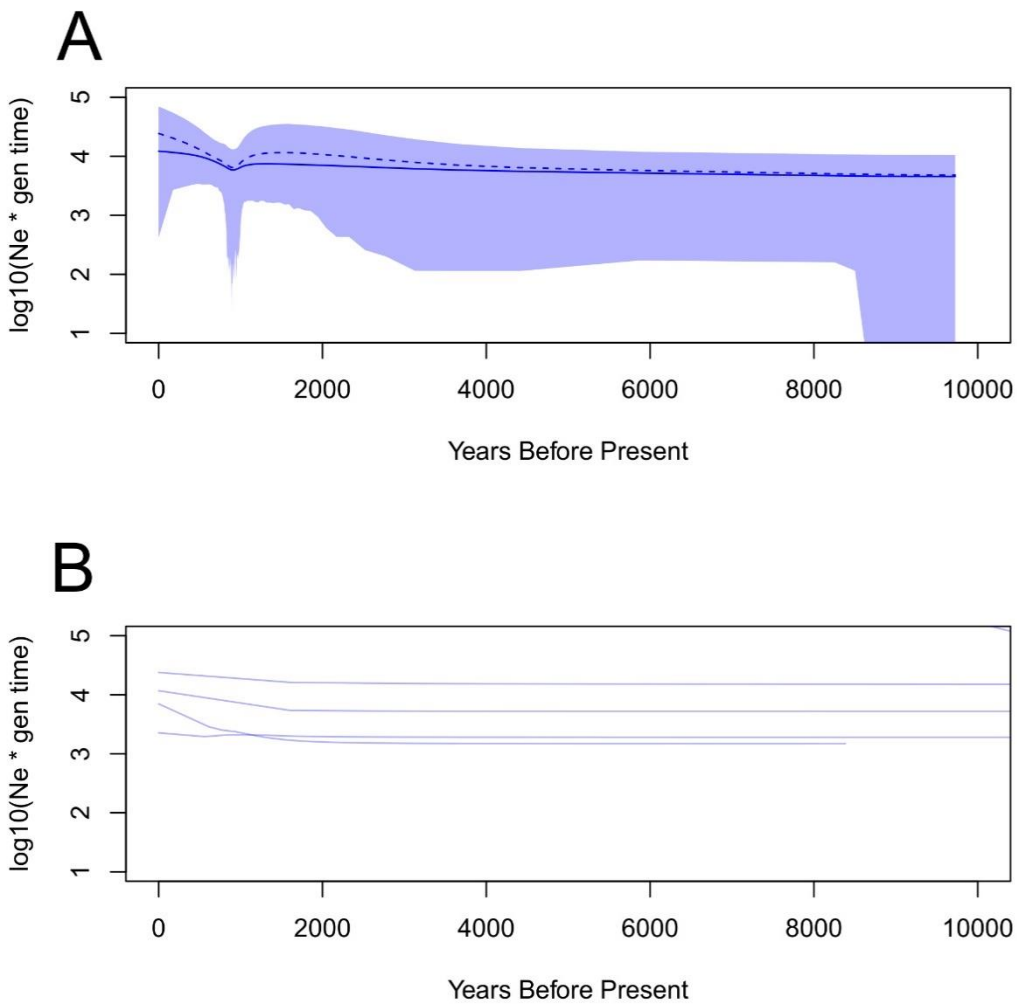
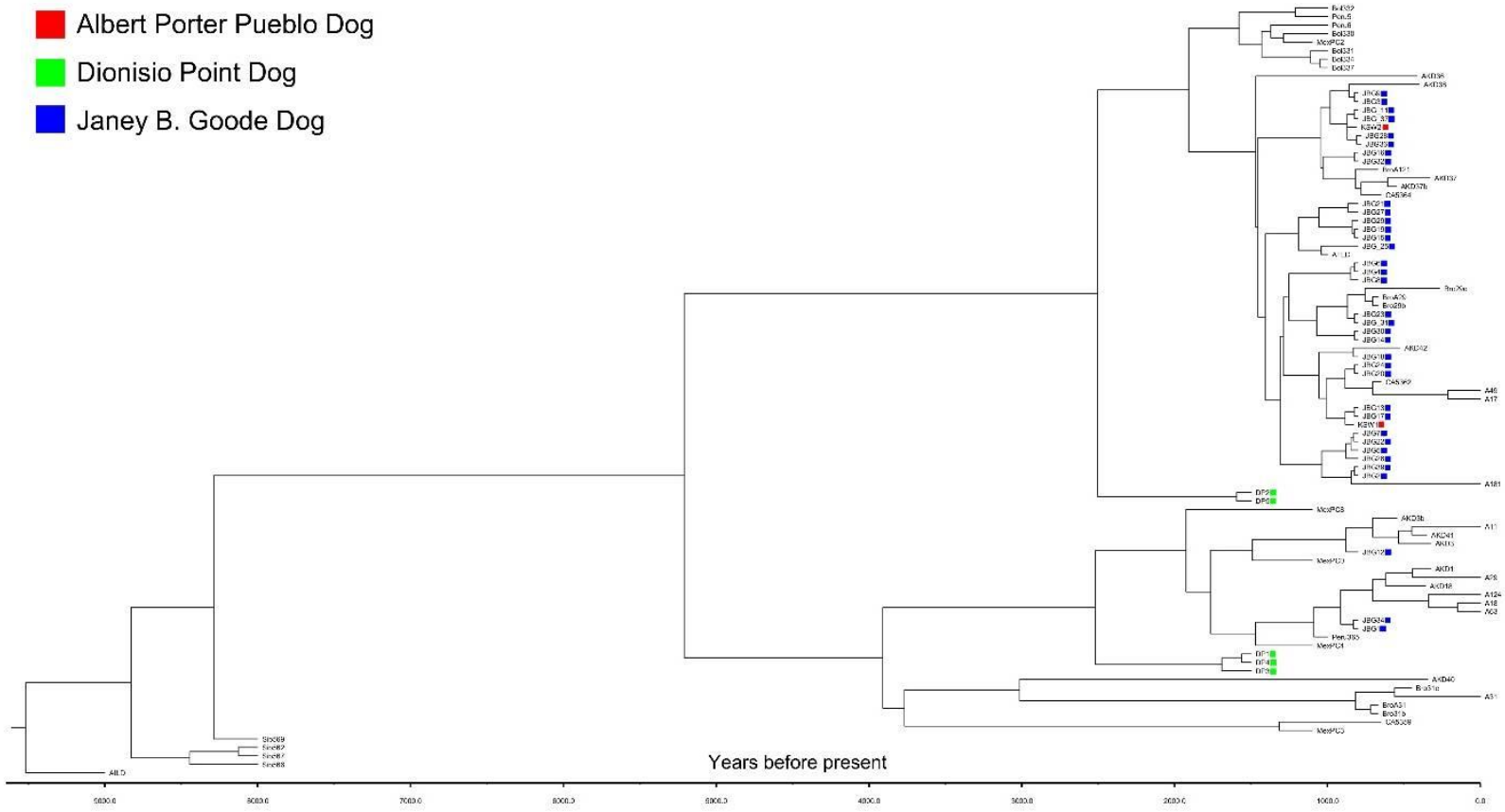


Figure 2.7: A phylogenetic tree generated from the extended Bayesian skyline plot analysis of the empirical dog data (ancient and modern). The sample includes all ancient dogs in the Americas that don't show wolf admixture as well as haplotypes that are shared with modern dogs. APP dogs are shown in red, DP dogs in green, and JBG dogs in blue. This phylogeny summarized the distribution of trees sampled during the MCMC analysis using the maximum clade credibility criterion with median branch lengths.



Tables

Table 2.1: A list of primers used in sequencing a portion of the hypervariable region of the mitochondrial genome of the individuals used in the study.

Amplicon	Forward	Reverse	Source
15421-15617	GCACCCAAAGCTGAGATTCT	GAGTTAATATGTCCTATGTAAGG	This study
15451-15744	TTCCCTGACACCCCTACATTCATATATT	GCCCGGAGCGAGAAGAGGGAC	Druzhkova 2013
15485-15694	CCCCTACTGTGCTATGTCAGTATCTCCAG	GGCATGGTGATTAAGCCCTTATTGGAC	Druzhkova 2013
15595-15796	CCTTACATAGGACATATTAACTC	AGAACCAGATGCCAGGTATAG	This study
15668-15863	GTCCAATAAGGGCTTAATCACCATGCC	TCCATCGAGATGTCCCATTTGCGA	Druzhkova 2013
15776-15982	CTATACCTGGCATCTGGTTCT	TTAGAGTTAGTGCCGTTGCG	This study
15963-16138	CGCAACGGCACTAACTCTAA	TACGTGTACCCTAAAACTATAT	This study

Table 2.2: Samples extracted at WSU, indicating amount of each element processed and the number of repeat silica extractions performed. Dionisio Point is abbreviated DP and Albert Porter Pueblo as APP.

Site	Sample ID	Element	amount extracted (mg)	# repeat silica
DP	DgRu-6 B1 SF10 PLM	Tooth-P3	28	0
DP	DgRu-6 B1 SF11 PLM	Tooth-Unerupted M3	35	0
DP	DgRu-6 B2 West 1 PLM	Tooth-M1	51	0
DP	DgRu-6 B2 West 2 PLM	Tooth-M2	32	0
DP	DgRu-6 B2 West 3 PLM	Mandible	55	2
DP	DgRu-6 B3 East 45	Tooth-P4	40	0
DP	DgRu-6 B3 E 58	Tooth-M2	47	0
DP	DgRu-6 B4 SE 2 PLM	Tooth-M2	31	0
APP	5MT123 PD 1305 FS42	Tooth-Unknown	54	0
APP	5MT123 PD 1936 FS2	Tooth-Incisor	42	1

Table 2.3: A list of mtDNA data of ancient dogs in the Americas in the literature. The general location of the archaeological site(s) is included, as well as the date (if the samples were dated) and the number of individuals.

Origin	Age	# individuals	Source
Siberia	8000 ybp	4	Losey 2013
Bolivia	>1000 ybp	5	Leonard 2002
Peru	1000 ybp	3	Leonard 2002
Argentina	1000 ybp	1	Thalmann 2013
Mexico	1400-800 ybp	5	Leonard 2002
Alaska	~400-600 ybp	11	Leonard 2002
Alaska	~200-800 ybp	7	Brown 2013
California	~900-400 ybp	3	Byrd 2013
Koster, Illinois	9000 ybp	1	Thalmann 2013
Florida	1000 ybp	1	Thalmann 2013
Western Canada	unknown	5	Koop 2000

Table 2.4: Haplotypes for each of the individuals from Janey B. Goode, Alfred Porter Pueblo and Dionisio Point, including a list of substitutions that differ from the reference dog mitochondrial genome (Genbank Accession NC002008). Haplotypes with an asterisk are novel and are identified as A191-A194. If multiple haplotypes are listed, the region of mtDNA that was sequenced for the individual is identical to multiple haplotypes. All sequences are available on Genbank. (accession numbers KJ189495-KJ189536).

Individual	Location	Date	Substitutions	Haplotype(s)
JBG 98-1	Janey B. Goode, IL	1000-1400 ybp	(none)	**A18, A19, A20**
JBG 457-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 34-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 34-2	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 457-2	Janey B. Goode, IL	1000-1400 ybp	15534.1C, 15627G, 15639A	**A136, A150**
JBG 635-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 726-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 741-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 786-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 1671-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 845-1	Janey B. Goode, IL	1000-1400 ybp	15639A	**A11, A89, A91, A123**
JBG 939-1	Janey B. Goode, IL	1000-1400 ybp	15484G, 15560.X, 15627G, 15639A	**A191***
JBG 975-1	Janey B. Goode, IL	1000-1400	15627G, 15639A	**A3, A4, A5, A6, A7, A8,**

Table 2.4 (cont'd)

		ybp		**A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 2255-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 2356-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 2601-1	Janey B. Goode, IL	1000-1400 ybp	15484G, 15627G, 15639A	**A192***
JBG 3222-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 1724-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 4109-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 4109-2	Janey B. Goode, IL	1000-1400 ybp	15461T, 15627G, 15639A	**A193***
JBG 4344-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 2793-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 2793-2	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 4939-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 5267-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62,**

Table 2.4 (cont'd)

				A130, A138, A140, A158, A182, A184
JBG 3134-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 5499-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 5267-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 5606-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 5609-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 5819-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 6134-1	Janey B. Goode, IL	1000-1400 ybp	(none)	**A18, A19, A20**
JBG 6963-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 7023-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
JBG 7458-1	Janey B. Goode, IL	1000-1400 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158, A182, A184**
5MT123 PD 1305 FS42	Alfred Porter Pueblo, CO	740-920 ybp	15484G, 15627G, 15639A	**A192***
5MT123 PD 1936 FS2	Alfred Porter Pueblo, CO	740-920 ybp	15627G, 15639A	**A3, A4, A5, A6, A7, A8, A9, A10, A22, A62, A130, A138, A140, A158,**

Table 2.4 (cont'd)

				A182, A184
DgRu-6 B1 SF11 PLM	Dionisio Point, BC	1500 ybp	15628C, 15639A	**A194***
DgRu-6 B2 West 1 PLM	Dionisio Point, BC	1500 ybp	15553G, 15627G, 15639A	**A102**
DgRu-6 B2 West 2 PLM	Dionisio Point, BC	1500 ybp	15628C, 15639A	**A194***
DgRu-6 B4 SE 2 PLM	Dionisio Point, BC	1500 ybp	15628C, 15639A	**A194***
DgRu-6 B3 East 45	Dionisio Point, BC	1500 ybp	15553G, 15627G, 15639A	**A102**

Table 2.5: Measures of genetic diversity of ancient dog populations by region. For each population, the number of individuals and haplotypes is recorded, as well as Theta S, a measure of the number of segregating sites weighted by sample size. Nucleotide diversity is also included, a measure of diversity that incorporates pairwise comparisons between samples, and was calculated both with equal haplotype frequencies and with weighted haplotype frequencies.

Population	# Individuals	# Haplotypes	Theta S	Nucleotide Diversity, weighted by haplotype frequency	Nucleotide diversity, not weighted by haplotype frequency
Janey B. Goode	34	7	1.223 (.636)	.00272 (.00243)	.0103 (.00728)
Dionisio Point	5	2	1.44 (1.016)	.00783 (.0063)	.0130 (.0151)
Siberia	4	3	1.091 (.876)	.00435 (.00431)	.00580 (.00596)
Alaska	18	11	6.105 (2.444)	.0198 (.0114)	.0208 (.0124)
Mexico	5	5	6.24 (3.446)	.0235 (.0159)	.0235 (.0159)
California	3	3	4.667 (3.127)	.0217 (.0180)	.0217 (.0180)
Western Canada	5	4	4.32 (2.484)	.0182 (.0126)	.0224 (.0163)
Bolivia	5	3	0.96 (0.758)	.00349 (.00348)	.00582 (.00598)
Peru	3	3	3.333 (2.323)	.0145 (.0126)	.0145 (.0126)
Wolf	7	6	5.128 (1.635)	.0231 (.0145)	.0237 (.0153)
Dogs	83	36	7.214 (2.135)	.0126 (.0007)	.0195 (.0110)

Table 2.6: Measures of genetic diversity of ancient dog populations by archaeological site. For each population, the number of individuals and haplotypes is recorded, as well as Theta S, a measure of the number of segregating sites weighted by sample size. Nucleotide diversity is also included, a measure of diversity that incorporates pairwise comparisons between samples, and was calculated both with equal haplotype frequencies and with weighted haplotype frequencies.

Population	# Individuals	# Haplotypes	Theta S	Nucleotide Diversity, weighted by haplotype frequency	Nucleotide Diversity, not weighted by haplotype frequency
Janey B. Goode	34	7	1.223 (.636)	.00272 (.00243)	.0103 (.00728)
Dionisio Point	5	2	1.44 (1.016)	.00783 (.0063)	.0130 (.0151)
Siberia	4	3	1.091 (.876)	.00435 (.00431)	.00580 (.00596)
Fairbanks, AK	11	9	3.265 (1.786)	.0199 (.0127)	.00177 (.0111)
W. Alaska	7	3	5.463 (2.479)	.0162 (.00999)	.0232 (.0191)
Tula, Mexico	3	3	7.333 (4.727)	.0318 (.0256)	.0318 (.0256)
California	3	3	4.667 (3.127)	.0217 (.0180)	.0217 (.0180)
Western Canada	5	4	4.32 (2.484)	.0182 (.0126)	.0224 (.0163)
Bolivia	3	3	0.96 (0.758)	.00349 (.00348)	.00582 (.00598)
Peru	3	3	3.333 (2.323)	.0145 (.0126)	.0145 (.0126)

References

Achilli, A., Perego, U. a, Bravi, C.M., Coble, M.D., Kong, Q.-P., Woodward, S.R., Salas, A., Torroni, A., Bandelt, H.-J., 2008. The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. PloS ONE. 3, e1764.

Achilli, A., Perego, U. a, Lancioni, H., Olivieri, A., Gandini, F., Hooshiar Kashani, B., Battaglia, V., Grugni, V., Angerhofer, N., Rogers, M.P., Herrera, R.J., Woodward, S.R., Labuda, D., Smith, D.G., Cybulski, J.S., Semino, O., Malhi, R.S., Torroni, A., 2013. Reconciling migration models to the Americas with the variation of North American native mitogenomes. Proceedings of the National Academy of Sciences of the United States of America. 110, 14308–13.

Ardalan, A., Kluetsch, C.F.C., Zhang, A.-B., Erdogan, M., Uhlén, M., Houshmand, M., Tepeli, C., Ashtiani, S.R.M., Savolainen, P., 2011. Comprehensive study of mtDNA among Southwest Asian dogs contradicts independent domestication of wolf, but implies dog-wolf hybridization. Ecology and Evolution. 1, 373–85.

Axelsson, E., Ratnakumar, A., Arendt, M.-L., Maqbool, K., Webster, M.T., Perloski, M., Liberg, O., Arnemo, J.M., Hedhammar, A., Lindblad-Toh, K., 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature. 495, 360-364.

Bandelt, H.-J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution. 16, 37–48.

Barta, J.L., 2006. Addressing issues of domestication and cultural continuity. Doctoral Dissertation.

Bielejec, F., Lemey, P., Carvalho, L., 2014. piBUSS: a parallel BEAST/BEAGLE utility for sequence simulation under complex evolutionary scenarios. BMC Bioinformatics. 15, 133.

Borgic, Q.L., Galloy, J.M., 2004. Domesticated Dog Remains from the Janey B. Goode Site. In Joint Meetings of the Southeastern Archaeological Conference and the Midwestern Archaeological Conference, St. Louis, Missouri, October (Vol. 2).

Boyko, A.R., Boyko, R.H., Boyko, C.M., Parker, H.G., Castelhano, M., Corey, L., Degenhardt, J.D., Auton, A., Hedimbi, M., Kityo, R., Ostrander, E. a, Schoenebeck, J., Todhunter, R.J., Jones, P., Bustamante, C.D., 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. Proceedings of the National Academy of Sciences of the United States of America. 106, 13903–8.

Brown, S.K., Darwent, C.M., Sacks, B.N., 2013. Ancient DNA evidence for genetic continuity in arctic dogs. Journal of Archaeological Science. 40, 1279–1288.

Byrd, B.F., Cornellas, A., Eerkens, J.W., Rosenthal, J.S., Carpenter, T.R., Leventhal, A., Leonard, J. a., 2013. The role of canids in ritual and domestic contexts: new ancient DNA insights from complex hunter–gatherer sites in prehistoric Central California. Journal of Archaeological Science. 40, 2176–2189.

Castroviejo-Fisher, S., Skoglund, P., Valadez, R., Vilà, C., Leonard, J. a, 2011. Vanishing native American dog lineages. BMC evolutionary biology. 11, 73.

Clutton-Brock, J., 1995. Origins of the dog: domestication and early history. In: Serpell, J. (Ed.), The Domestic Dog: Its Evolution, Behaviour and Interactions with People. Cambridge University Press, Cambridge, pp. 7–20.

Coppinger, R., Coppinger, L., 2001. Dogs: A New Understanding of Canine Origin, Behavior, and Evolution. University of Chicago Press, Chicago.

Cui, Y., Lindo, J., Hughes, C.E., Johnson, J.W., Hernandez, A.G., Kemp, B.M., Ma, J., Cunningham, R., Petzelt, B., Mitchell, J., Archer, D., Cybulski, J.S., Malhi, R.S., 2013. Ancient DNA analysis of mid-holocene individuals from the Northwest Coast of North America reveals different evolutionary paths for mitogenomes. PloS ONE. 8, e66948.

Dillehay, T., Collins, M., 1988. Early cultural evidence from Monte Verde in Chile. Nature. 332, 150–152.

Ding, Z.-L., Oskarsson, M., Ardalan, A., Angleby, H., Dahlgren, L.-G., Tepeli, C., Kirkness, E., Savolainen, P., Zhang, Y.-P., 2012. Origins of domestic dog in southern East Asia is supported by analysis of Y-chromosome DNA. Heredity. 108, 507–14.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology. 4, e88.

Drummond, A.J., Suchard, M. A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 29, 1969–73.

Druzhkova, A.S., Thalmann, O., Trifonov, V. a., Leonard, J. a., Vorobieva, N. V., Ovodov, N.D., Graphodatsky, A.S., Wayne, R.K., 2013. Ancient DNA Analysis Affirms the Canid from Altai as a Primitive Dog. PLoS ONE. 8, e57754.

Erlandson, J., 2007. The Kelp Highway Hypothesis: Marine Ecology, the Coastal Migration Theory, and the Peopling of the Americas. The Journal of Island and Coastal Archaeology. 2, 161–174.

Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10, 564–567.

Fagundes, N.J.R., Kanitz, R., Eckert, R., Valls, A.C.S., Bogo, M.R., Salzano, F.M., Smith, D.G., Jr, W.A.S., Zago, M.A., Ribeiro-dos-santos, A.K., Santos, S.E.B., Petzl-erler, M.L., Bonatto, S.L., 2008. Mitochondrial Population Genomics Supports a Single Pre-Clovis Origin with a Coastal Route for the Peopling of the Americas. American Journal of Human Genetics. 82, 583–592.

Ferreira, M., Suchard, M., 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. Canadian Journal of Statistics. 36, 355–368.

Forster, P., Harding, R., Torroni, a, Bandelt, H.J., 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. American Journal of Human Genetics. 59, 935–45.

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E. a., Marques-Bonet, T., Wayne, R.K.,

Novembre, J., 2014. Genome Sequencing Highlights the Dynamic Early History of Dogs. PLoS Genetics. 10, e1004016.

Fuller, T.K., 1995. Comparative population dynamics of North American wolves and African Wild Dogs. In: Carbyn, L.N., Fritts, S.H., Seip, D.R. (Eds.), Ecology and Conservation of Wolves in a Changing World. Canadian Circumpolar Institute, Edmonton, Canada.

Galloy, J.M., 2010. PART V : WOODLAND PERIOD The Janey B . Goode Site (11S1232): Highlights of Investigations at a Massive Late Prehistoric Site in the American Bottom. Illinois Archaeology. 22, 529–552.

Germonpré, M., Sablin, M. V., Stevens, R.E., Hedges, R.E.M., Hofreiter, M., Stiller, M., Després, V.R., 2009. Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine and Russia: osteometry, ancient DNA and stable isotopes. Journal of Archaeological Science. 36, 473–490.

Gilbert, M.T.P., Hansen, A.J., Willerslev, E., Rudbeck, L., Barnes, I., Lynnerup, N., Cooper, A., 2003. Characterization of genetic miscoding lesions caused by postmortem damage. American Journal of Human Genetics. 72, 48–61.

Grier, C., 2006. Temporality in Northwest Coast Households. In: Sobel, E.A., Gahr, D.A.T., Ames, K.M. (Eds.), Household Archaeology on the Northwest Coast. International Monographs in Prehistory, Ann Arbor, pp. 97–119.

Grier, C., Flanigan, K., Winters, M., Jordan, L.G., Lukowski, S., Kemp, B.M., 2013. Using ancient DNA identification and osteometric measures of archaeological Pacific salmon vertebrae for reconstructing salmon fisheries and site seasonality at Dionisio Point, British Columbia. Journal of Archaeological Science. 40, 544–555.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evolutionary Biology. 8, 289.

Hoffecker, J., Powers, W., Goebel, T., 1993. The colonization of Beringia and the peopling of the New World. Science. 259, 46–53.

Hooshiar Kashani, B., Perego, U. A., Olivieri, A., Angerhofer, N., Gandini, F., Carossa, V., Lancioni, H., Semino, O., Woodward, S.R., Achilli, A., Torroni, A., 2012. Mitochondrial haplogroup C4c: a rare lineage entering America through the ice-free corridor? American Journal of Physical Anthropology. 147, 35–9.

Kemp, B., Malhi, R., McDonough, J., Bolnick, D.A., Eshleman, J.A., Rickards, O., Martinez-Labarga, C., Johnson, J.R., Lorenz, J.G., Dixon, E.J., Fifield, T.E., Heaton, T.H., Worl, R., Smith, D.G., 2007. Genetic Analysis of Early Holocene Skeletal Remains From Alaska and its Implications for the Settlement of the Americas. American Journal of Physical Anthropology 132, 605–621.

Kemp, B.M., Monroe, C., Judd, K.G., Reams, E., Grier, C., 2014. Evaluation of methods that subdue the effects of polymerase chain reaction inhibitors in the study of ancient and degraded DNA. Journal of Archaeological Science. 42, 373–380.

Kemp, B.M., Schurr, T.G., 2010. Ancient and modern genetic variation in the Americas. In: Auerbach, B.

(Ed.), Human Variation in the Americas: The Integration of Archeology and Biological Anthropology. Carbondale, IL, pp. 12–50.

Kitchen, A., Miyamoto, M.M., Mulligan, C.J., 2008. A three-stage colonization model for the peopling of the Americas. PloS ONE. 3, e1596.

Klütsch, C.F.C., de Caprona, M.D.C., 2010. The IGF1 small dog haplotype is derived from Middle Eastern grey wolves: a closer look at statistics, sampling, and the alleged Middle Eastern origin of small dogs. BMC Biology. 8, 119.

Koop, B.F., Burbidge, M., Byun, A., Rink, U., Crockford, S., 2000. Ancient DNA evidence of a separate origin for North American indigenous dogs. In: Crockford, S.J. (Ed.), Dogs Through Time: An Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. Archaeopress.

Kukekova, A. V, Temnykh, S. V, Johnson, J.L., Trut, L.N., Acland, G.M., 2012. Genetics of behavior in the silver fox. Mammalian Genome. 23, 164–77.

Leonard, J. a, Wayne, R.K., Wheeler, J., Valadez, R., Guillén, S., Vilà, C., 2002. Ancient DNA evidence for Old World origin of New World dogs. Science. 298, 1613–6.

Li, Y., Vonholdt, B.M., Reynolds, A., Boyko, A.R., Wayne, R.K., Wu, D.-D., Zhang, Y.-P., 2013. Artificial selection on brain expressed genes during the domestication of dog. Molecular Biology and Evolution. 30, 1867–1876.

Li, Y., Wu, D.-D., Boyko, A.R., Wang, G.-D., Wu, S.-F., Irwin, D.M., Zhang, Y.-P., 2014. Population Variation Revealed High-Altitude Adaptation of Tibetan Mastiffs. Molecular Biology and Evolution. 31, 1200–1205.

Losey, R.J., Garvie-Lok, S., Leonard, J. a, Katzenberg, M.A., Germonpré, M., Nomokonova, T., Sablin, M. V, Goriunova, O.I., Berdnikova, N.E., Savel'ev, N. a, 2013. Burying dogs in ancient cis-baikal, siberia: temporal trends and relationships with human diet and subsistence practices. PloS ONE. 8, e63740.

Malhi, R.S., Cybulski, J.S., Tito, R.Y., Johnson, J., Harry, H., Dan, C., 2010. Brief communication: mitochondrial haplotype C4c confirmed as a founding genome in the Americas. American Journal of Physical Anthropology. 141, 494–7.

Malhi, R.S., Eshleman, J. a, Greenberg, J. a, Weiss, D. a, Schultz Shook, B. a, Kaestle, F. a, Lorenz, J.G., Kemp, B.M., Johnson, J.R., Smith, D.G., 2002. The structure of diversity within New World mitochondrial DNA haplogroups: implications for the prehistory of North America. American Journal of Human Genetics. 70, 905–19.

Malhi, R.S., Kemp, B.M., Eshleman, J. a., Cybulski, J., Smith, D.G., Cousins, S., Harry, H., 2007. Mitochondrial haplogroup M discovered in prehistoric North Americans. Journal of Archaeological Science. 34, 642–648.

Malhi, R.S., Smith, D.G., 2002. Brief communication: Haplogroup X confirmed in prehistoric North America. American Journal of Physical Anthropology. 119, 84–6.

Matisoo-Smith, E., Robins, J.H., 2004. Origins and dispersals of Pacific peoples: evidence from mtDNA phylogenies of the Pacific rat. Proceedings of the National Academy of Sciences of the United States of America. 101, 9167–72.

Meltzer, D., 2010. First Peoples in a New World: Colonizing Ice Age America. University of California Press.

Morey, D.F., Wiant, M.D., 1992. Early Holocene Domestic Dog Burials from the North American Midwest. Current Anthropology. 33, 224–230.

Mulligan, C.J., Kitchen, A., Miyamoto, M.M., 2008. Updated three-stage model for the peopling of the Americas. PloS ONE. 3, e3199.

O'Fallon, B.D., Fehren-Schmitz, L., 2011. Native Americans experienced a strong population bottleneck coincident with European contact. Proceedings of the National Academy of Sciences of the United States of America. 108, 20444–8.

Okumura, N., Ishiguro, N., Nakano, M., Matsui, A., Shigehara, N., Nishimoto, T., Sahara, M., 1999. Variations in Mitochondrial DNA of Dogs Isolated from Archaeological Sites in Japan and Neighboring Islands. Anthropological Science. 107, 213–228.

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A., Luo, L.-Y., Angleby, H., Ardalan, A., Ekström, C., Sköllermo, A., Lundeberg, J., Matsumura, S., Leitner, T., Zhang, Y.-P., Savolainen, P., 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. Molecular Biology and Evolution. 26, 2849–64.

Perego, U. A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., Hooshiar Kashani, B., Ritchie, K.H., Scozzari, R., Kong, Q.-P., Myres, N.M., Salas, A., Semino, O., Bandelt, H.-J., Woodward, S.R., Torroni, A., 2009. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Current Biology. 19, 1–8.

Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F. a, Mountain, J.L., Misra, R., Carter, N.P., Lee, C., Stone, A.C., 2007. Diet and the evolution of human amylase gene copy number variation. Nature Genetics. 39, 1256–60.

Rambaut, A., Drummond, A.J., 2007. Tracer v. 1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir, V., Yadav, R., Malaspinas, A.-S., White, S.S., Allentoft, M.E., Cornejo, O.E., Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M., Korneliussen, T.S., Meltzer, D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S., Brunak, S., Sicheritz-Ponten, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica, A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., Willerslev, E., 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature. 506, 225–9.

Ryan, S.C., 2004. Albert Porter Pueblo (Site 5MT123), Montezuma County, Colorado: Annual Report, 2004 Field Season.

Saetre, P., Lindberg, J., Leonard, J. a, Olsson, K., Pettersson, U., Ellegren, H., Bergström, T.F., Vilà, C., Jazin, E., 2004. From wild wolf to domestic dog: gene expression changes in the brain. Brain research. Molecular Brain Research. 126, 198–206.

Savolainen, P., Zhang, Y., Luo, J., Lundeberg, J., Leitner, T., 2002. Genetic evidence for an East Asian origin of domestic dogs. Science. 298, 1610–3.

Scheinfeldt, L.B., Soi, S., Thompson, S., Ranciaro, A., Woldemeskel, D., Beggs, W., Lambert, C., Jarvis, J.P., Abate, D., Belay, G., Tishkoff, S. a, 2012. Genetic adaptation to high altitude in the Ethiopian highlands. Genome Biology. 13, R1.

Schroeder, K.B., Jakobsson, M., Crawford, M.H., Schurr, T.G., Boca, S.M., Conrad, D.F., Tito, R.Y., Osipova, L.P., Tarskaia, L. a, Zhadanov, S.I., Wall, J.D., Pritchard, J.K., Malhi, R.S., Smith, D.G., Rosenberg, N. a, 2009. Haplotypic background of a private allele at high frequency in the Americas. Molecular Biology and Evolution. 26, 995–1016.

Schurr, T.G., Sherry, S.T., 2004. Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. American Journal of Human Biology. 16, 420–39.

Schwartz, M., 1997. A History of Dogs in the Early Americas. Yale University Press, New Haven, CT.

Smith, D.G., Malhi, R.S., Eshleman, J., Lorenz, J.G., Kaestle, F. a, 1999. Distribution of mtDNA haplogroup X among Native North Americans. American Journal of Physical Anthropology. 110, 271–84.

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V., Richards, M.B., 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. American Journal of Human Genetics. 84, 740–59.

Takahasi, S., Miyahara, K., 2002. Lineage classification of canine inheritable disorders using mitochondrial DNA haplotypes. The Journal of Veterinary Medical Science. 64, 255-259.

Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D.G., Mulligan, C.J., Bravi, C.M., Rickards, O., Martinez-Labarga, C., Khusnutdinova, E.K., Fedorova, S. a, Golubenko, M. V, Stepanov, V. a, Gubina, M. a, Zhadanov, S.I., Ossipova, L.P., Damba, L., Voevoda, M.I., Dipierri, J.E., Villems, R., Malhi, R.S., 2007. Beringian standstill and spread of Native American founders. PloS ONE. 2, e829.

Tamura, K., Peterson, D., Peterson, N., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and Evolution. 28, 2731-9.

Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpre, M.., Sablin, M.V., Lopez-Giraldez, F., Domingo-Roura, X., Napierala, H., Uerpmann, H.-P., Loponte, D.M., Acosta, A.., Giemsch, L., Schmitz, R.W., Worthington, B., Buikstra, J.E., Druzhkova, A., Graphodatsky, A.S., Ovodov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfli, K.-P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., Wayne, R.K., 2013. Complete Mitochondrial Genomes of Ancient Canids Suggest a European Origin of Domestic Dogs. Science. 342, 871–874.

Torroni, a, Schurr, T.G., Cabell, M.F., Brown, M.D., Neel, J. V, Larsen, M., Smith, D.G., Vullo, C.M., Wallace, D.C., 1993a. Asian affinities and continental radiation of the four founding Native American mtDNAs. American Journal of Human Genetics. 53, 563–90.

Torroni, a, Sukernik, R.I., Schurr, T.G., Starikorskaya, Y.B., Cabell, M.F., Crawford, M.H., Comuzzie, a G., Wallace, D.C., 1993b. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. American Journal of Human Genetics. 53, 591–608.

Tsuda, K., Kikkawa, Y., Yonekawa, H., 1997. Extensive interbreeding occurred among multiple matriarchal ancestors during the domestication of dogs : Evidence from inter- and intraspecies polymorphisms in the D-loop region of mitochondrial DNA between dogs and wolves. Genes and Genetic Systems. 72, 229–238.

van Asch, B., Zhang, A.-B., Oskarsson, M.C.R., Klütsch, C.F.C., Amorim, A., Savolainen, P., 2013. Pre-Columbian origins of Native American dog breeds, with only limited replacement by European dogs, confirmed by mtDNA analysis. Proceedings of the Royal Society B: Biological Sciences. 280, 20131142.

Vila, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J., Wayne, R.K., 1997. Multiple and Ancient Origins of the Domestic Dog. Science. 276, 1687–1689.

Villanea, F. A., Bolnick, D. A., Monroe, C., Worl, R., Cambra, R., Leventhal, A., Kemp, B.M., 2013. Brief communication: Evolution of a specific O allele (O1v(G542A)) supports unique ancestry of Native Americans. American Journal of Physical Anthropology. 151, 649–57.

von Valkenburgh, B., 1990. Skeletal and dental predictors of body mass in carnivores. In: Body Size in Mammalian Paleobiology; Estimation and Biological Implications. pp. 181–205.

Vonholdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E., Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D. a, Auton, A., Reynolds, A., Bryc, K., Brisbin, A., Knowles, J.C., Mosher, D.S., Spady, T.C., Elkahloun, A., Geffen, E., Pilot, M., Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.-L., Zhang, Y.-P., Bustamante, C.D., Ostrander, E. a, Novembre, J., Wayne, R.K., 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature. 464, 898–902.

Walberg, M., Clayton, D., 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Research. 9, 5411–5421.

Wang, S., Lewis, C.M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M. V, Molina, J. a, Gallo, C., Mazzotti, G., Poletti, G., Hill, K., Hurtado, A.M., Labuda, D., Klitz, W., Barrantes, R., Bortolini, M.C., Salzano, F.M., Petzl-Erler, M.L., Tsuneto, L.T., Llop, E., Rothhammer, F., Excoffier, L., Feldman, M.W., Rosenberg, N. a, Ruiz-Linares, A., 2007. Genetic variation and population structure in native Americans. PLoS Genetics. 3, e185.

Watterson, G., 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology. 276, 256–276.

Xu, S., Li, S., Yang, Y., Tan, J., Lou, H., Jin, W., Yang, L., Pan, X., Wang, J., Shen, Y., Wu, B., Wang, H., Jin, L., 2011. A genome-wide search for signals of high-altitude adaptation in Tibetans. Molecular Biology and Evolution. 28, 1003–11.

Yohe, R., Pavesic, M., 2000. Early archaic domestic dogs from Western Idaho, USA. In: Dogs Through Time: An Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. pp. 93–104.

CHAPTER THREE: MITOCHONDRIAL GENOME SEQUENCING OF ANCIENT DOGS IN THE AMERICAS TO UNDERSTAND THEIR DEMOGRAPHIC HISTORY

Abstract

Several ancient DNA studies have been conducted on dogs in the Americas, yet all but two have sequenced only the hypervariable region of the mitochondrial genome. In this study, we sequenced 68 complete mitochondrial genomes (mitogenomes) of ancient dogs from 19 archaeological sites, ranging in age from 9000 to 800 years before present (ybp) to gain insight into the population history of dogs in the Americas. We measured the genetic diversity of each population sampled, compared the ancient dogs to modern dogs and wolves, and modeled demographic changes in the population over time. Contrary to previous research, we found that no ancient American haplotypes are shared with modern dog haplotypes, although the closest modern dog haplogroup is A4, which is found in some Asian breeds and Chihuahuas. Additionally, consistent with some previous research, we found that dogs likely arrived in the Americas with the initial human entry into the Americas. We identified similarities between dog and human demographic histories, specifically regarding connections between Southeast and Midwest populations, and populations along the Pacific coast. We also found that an increase in dog population diversity in the Midwest coincides with the Woodland-Mississippian cultural transition and hypothesized movement of other populations into the Midwest during the Mississippian period. Sequencing complete mitochondrial genomes provides a clearer picture of dog population history in the Americas, as well as the connections between dog and human populations.

Introduction

Domestic and pest species have lived and traveled with humans across vast regions of the world. Therefore, genetic patterns of these species associated with humans can be used to infer human migrations. For example, the demographic history of chickens (Thomson et al., 2014), pigs (Larson et al., 2007), and rats (Matisoo-Smith and Robins, 2004) in Oceania have been used to reconstruct the migration routes of Polynesians. Mice have traveled worldwide on ships, and their demographic history reveals the voyages of the Vikings and other ancient mariners (Jones, Eager, Gabriel, Jóhannesdóttir, & Searle, 2013). Dogs have lived with humans for millennia, and their long history with humans enables them to be useful as biological proxies. Dogs have also been used to clarify the history of humans, both in New Zealand (Greig et al., 2015) and in the Arctic (Brown et al., 2013). They have even been used as a dietary proxy for humans, through the analysis of stable isotopes (Noe-Nygaard, 1988; Clutton-brock and Noe-nygaardb, 1990; White et al., 2001; Allitt et al., 2009; Rick et al., 2011; Guiry, 2012).

Dog domestication occurred approximately 15,000 - 20,000 years before present (ybp) (Pang et al., 2009; Vonholdt et al., 2010; Thalmann et al., 2013; Freedman et al., 2014). Some ancient canid remains that date to at least 30,000 ybp and resemble dogs have been identified (Ovodov et al., 2011), but it is difficult to distinguish between dogs and wolves based on morphology alone (Drake et al., 2015), and it is likely that if there were early domestication attempts, that these lineages did not survive to contribute to modern dog populations (Druzhkova et al., 2013). The origin of dogs has been studied thoroughly, but the exact location of dog domestication remains controversial. Regions including Europe (Thalmann et al., 2013; Frantz et al., 2016), the Middle East (Vonholdt et al., 2010), Southeast Asia (Pang et al., 2009; Ding et al., 2012), Africa (Boyko et al., 2009), and Central Asia (Shannon et al., 2015) have all been suggested. It was long thought that dog domestication occurred only once (Pang et al., 2009), but recent study of ancient dogs in Eurasia has suggested that dogs were domesticated independently in Asia and in Europe, and that the Asian population expanded westward and eventually replaced the early European dogs

(Frantz et al., 2016).

Using present-day dog genetic diversity to infer the history of the domestic dog is difficult because many of today's dog breeds were only formed in the past three hundred years, and large population bottlenecks occurred in the creation of those breeds, resulting in very low genetic diversity in modern dogs (Karlsson et al., 2007; Larson et al., 2012). Therefore, by only using modern breed dogs, one is limited in the conclusions one can make about the historically large and diverse dog population. Some studies incorporate "village dogs", dogs of indeterminate breed found all over the world that likely did not experience the same bottlenecks, to overcome this problem (Boyko et al., 2009; Brown et al., 2011; Shannon et al., 2015). Another approach is to use DNA from ancient dogs, which pre-date the bottlenecks that were part of breed formation (Leonard et al., 2002; Verginelli et al., 2005; Thalmann et al., 2013; Witt et al., 2015; Frantz et al., 2016). Village dogs show much higher levels of genetic diversity than breed dogs (Boyko et al., 2009; Shannon et al., 2015), and the coalescence dates of these populations are more likely to reflect the timing of dog domestication, not breed formation. Studies of dogs from the Americas (Leonard et al., 2002; Castroviejo-Fisher et al., 2011; van Asch et al., 2013) and Europe (Deguilloux et al., 2009; Frantz et al., 2016) have shown a shift in mitochondrial haplotype frequencies over time, suggesting that the original dog population in these areas was replaced by the ancestors of modern breed dogs (Sacks et al., 2013). Therefore, to infer the early history of the domestic dog, it is necessary to study ancient dogs.

There have been multiple genetic studies of ancient dogs in the Americas, although all of these studies have been limited to mitochondrial DNA sequences (Leonard et al., 2002; Brown et al., 2013; Thalmann et al., 2013; Witt et al., 2015). Mitochondrial DNA is favorable for analysis in ancient DNA studies, because it has a high copy number and is more likely to be preserved in ancient samples than chromosomal DNA (Hagelberg and Clegg, 1991). Additionally, it is strictly maternally inherited and does not recombine (R E Giles et al., 1980), so it is an excellent marker to use in studies of ancestry as related haplotypes are easier to infer relative to the nuclear genome. The first study of ancient American dogs

showed that they likely traveled with humans from Eurasia, rather than being domesticated separately from North American gray wolves (Leonard et al., 2002). Comparison of ancient dogs to modern dogs shows that much of the mitochondrial genetic diversity in the ancient dogs has been lost (Castroviejo-Fisher et al., 2011), except in more remote locations like the Arctic (Brown et al., 2011), and in certain breeds, including the Chihuahua and the Carolina Feral Dog (van Asch et al., 2013). These findings suggest that to use genetic diversity to infer the history of dogs in the Americas, it is necessary to focus on ancient individuals. A study of ancient dogs across the Americas showed that different populations varied in genetic diversity, and that some mitochondrial haplotypes were widespread while others were unique to a particular archaeological site (Witt et al., 2015). It also suggests that dogs may have arrived during a later human migration to the Americas, close to 10,000 ybp.

Humans initially arrived to the Americas between 15,000-20,000 ybp by crossing Beringia from Asia (Kemp and Schurr, 2010; Meltzer and Holliday, 2010; Llamas et al., 2016). The source population for Native Americans likely derives from Siberia, and the Altai region has been implicated in multiple studies (Zegura et al., 2004; Dulik et al., 2012). The first migrants likely took a coastal route (Wang et al., 2007; Bodner et al., 2012; Achilli et al., 2013), and signs of human habitation have been found as far south as the Monte Verde site in Chile that date to over 14,000 ybp (Dillehay and Collins, 1988). Other indicators of human presence in the Americas during this time include evidence for mammoth butchering at the Schaefer-Hebior site in Wisconsin (Overstreet and Kolb, 2003), and human coprolites from Paisley Caves in Oregon (Jenkins et al., 2012). Inland expansion occurred once glaciation withdrew, and migration was likely halted for a significant period of time when the Bering land bridge was submerged roughly 10,000 – 11,0000 ybp (Hoffecker et al., 1993).

One drawback to previous mitochondrial DNA studies of ancient dogs is that all but one of them (Thalmann et al., 2013) examined a short region of the mitochondrial genome (mitogenome), known as the hypervariable region (HVR). The 300-base pair stretch of the genome has a higher substitution rate

than the rest of the mitogenome (Brown et al., 1979), and so it is often used in ancient DNA studies because it is a short region that harbors a lot of genetic diversity. However, the hypervariable region shows only a limited view of mitochondrial DNA diversity (Malhi et al., 2002; Pang et al., 2009; Duchêne et al., 2011; Gómez-Carballa et al., 2012). Although the HVR contains a lot of genetic variation relative to other genomic regions of the same size, sequencing such a small portion of the mitogenome masks additional variation in dog populations (Duchêne et al., 2011). Due to the rapid mutation rate of the HVR, homoplasies can occur in multiple lineages. This could make two unrelated individuals appear to share a more recent common ancestor than they actually do (Stoneking, 2000; Torres et al., 2006). Another concern with the HVR is that "back-mutations" can occur, in which a nucleotide that differs from the ancestral sequence can have a second mutation that makes it identical to the ancestral sequence, thus masking variation (Torres et al., 2006). This can be a concern with any DNA sequence, but occurs more commonly in the HVR due to its rapid mutation rate. Sequencing the complete mitogenome will show a more complete picture of genetic diversity in dogs. To illustrate this, Figure 3.1 shows two haplotype networks of the same individuals, with A showing the network of HVR sequences and B showing the network of mitogenome sequences. One "haplotype" in the HVR network resolves into multiple mitogenome haplotypes, and the mitogenome network also shows greater divergence between sequences, demonstrating the importance of sequencing the complete mitochondrial genome. However, only three mitogenomes have been sequenced for dogs in the Americas, limiting the analyses of dog demographic history one can perform (Thalmann et al., 2013).

By sequencing complete mitogenomes, it is possible to reconstruct a more detailed perspective of dog demographic history. It also enables us to re-examine the findings of previous studies of dogs in the Americas that utilized only the hypervariable region of the mitogenome. Specifically, we can determine if there are haplotypes shared between ancient and modern dogs, as discussed in Castroviejo-Fisher et al. (2011) and van Asch et al. (2013), and if dogs arrived to the Americas around 10,000 ybp, after

humans initially arrived to the Americas (Witt et al., 2015). We can also compare the demographic histories of ancient American dogs and Native Americans, and infer how dogs moved with humans and were affected by human cultural changes.

In this study, we sequenced 68 additional mitogenomes of ancient dogs in the Americas, to gain a better understanding of their demographic history, and to compare the history of dogs to human history in the Americas. We aimed to test two hypotheses about dog history in the Americas that have been supported by HVR data. First, we wanted to test the hypothesis that ancient American dog haplotypes are found in certain modern dog breeds (van Asch et al., 2013). We found that ancient American dog haplotypes are not found in modern dogs, suggesting that the ancient American dog population was replaced with European dogs. Second, we wanted to test the hypothesis that dogs arrived with humans to the Americas, rather than being domesticated separately in North America (Leonard et al., 2002). We found that the ancient wolves that were most closely related to ancient American dogs were from Siberia and Switzerland, supporting this hypothesis. We also identified two major clades of dogs in the Americas, with a deep divergence time corresponding to the timing of the entry of humans into the Americas. And finally, we found that the geographic patterns of genetic diversity in dogs correspond to similar geographic patterns in ancient Native Americans.

Methods

Sampling

A total of 77 individuals were sampled for sequencing, from 22 archaeological sites. A full list of the sites, the number of individuals represented, and their radiocarbon age if known, is provided in Table 3.1. A map of the Americas showing the location of each archaeological site is provided in Figure 3.2. Of these 77, 71 samples produced high-quality sequence data(at least 7x depth and coverage of over 15,000 bp), and were used in this analysis. This sample includes the three previously published sequences from Thalmann et al. (2013), and adds 68 new mitogenomes from North America and Siberia, with a range in

age from 1000 ybp to 9000 ybp.

DNA Extraction, Library Construction and Sequencing

Given the age of the dogs from which samples were taken, the DNA is degraded, damaged, and easily contaminated with DNA from living individuals (Willerslev and Cooper, 2005; Gilbert et al., 2007). To limit the risk of contamination, all DNA extraction methods were performed in an ancient DNA laboratory, following standard protocols when working with ancient DNA (Cooper and Poinar, 2000; Poinar, 2003; Kemp and Smith, 2005; Kemp and Smith, 2010). These protocols include wearing full-body suits, cleaning all surfaces regularly with bleach, and using a UV crosslinker to prevent contaminating DNA from interfering with sequencing DNA from these ancient individuals. One individual was extracted and sequenced twice, to confirm that the mitogenome sequence could be reproduced. Most of the DNA extractions were performed at the University of Illinois at Urbana Champaign, at the Carl R. Woese Institute for Genomic Biology, with methods described in Witt et al. (2015), but a subset of extractions was performed at the Centre for GeoGenetics at the University of Copenhagen, with methods described in Allentoft et al. (2015).

At the University of Illinois, genomic libraries were built for the extracts using the NEBNext DNA Library Prep Kit for Illumina (New England Biolabs). They were amplified twice, first using the NEBNext DNA polymerase and the associated index primers, to allow the samples to be pooled and sequenced together. The first amplification followed manufacturer instructions, and was repeated for twelve cycles. For the second amplification, Phusion High-Fidelity PCR Master Mix with HF buffer (New England Biolabs) was used, and four PCR reactions were made for each sample. Five uL of PCR product from the first amplification was added to each reaction, and the DNA was amplified according to manufacturer's instructions for 12 cycles. The four reactions for each sample were pooled and cleaned using Ampure XP (Beckman Coulter) and MagSi-DNA NGS^{PREP} beads (MagnaMedics), using an 80% ethanol: sample ratio. The library was visually examined on an agarose gel, and quantitated using a Qubit 1.0 fluorometer

(Thermo Fisher Scientific). Only libraries with concentrations of at least 20 ng/uL were used for capture.

To enrich for mitochondrial DNA, we developed a custom set of RNA baits as part of a MYbaits kit (Mycroarray) that covered the complete dog mitogenome with 4x tiling density. Captures were performed using the Mycroarray manual version 3.01 at the University of Illinois, with a 60° C incubation for 28 hours. The heat elution step was skipped, and the capture was amplified using KAPA Hi-Fi polymerase, following manufacturer's instructions for 16 cycles. The PCR reaction was cleaned using MagSi-DNA NGS^{PREP} beads (MagnaMedics), and the capture was visually examined on an agarose gel and quantitated using a Qubit 1.0 Fluorimeter, following the manufacturer's instructions. If the DNA concentration of the amplified capture was lower than 20 ng/uL, the capture was reamplified for 8 cycles using the KAPA polymerase prior to sequencing.

At the University of Copenhagen, genomic libraries were built from the extracts using the NEBNext DNA Library Master Mix Set 2 (New England Biolabs), with modifications. The End Repair mix was incubated for 20 minutes at 12° C and 15 minutes at 37° C. The Quick Ligation mix was incubated for 20 minutes at 20° C. The Fill-In mix was incubated for 20 minutes at 65° C and 20 minutes at 80° C. Each step was purified using a Qiagen MinElute PCR Purification Kit. The protocol was followed as directed except that a differing amount of EB Buffer was used for each mix (30 uL for End Repair, and 42 uL for Quick Ligation) and the column was incubated for 15 minutes at 37° C prior to elution. The finished libraries were amplified using Taq Gold in a mix that included 10 uL Taq Gold Buffer, 2.5 mM $MgCl_2$, 0.8 mM uL BSA, 0.08 mM dNTPs, 0.2 µM of each of Illumina's Multiplexing PCR primer and a custom-designed index primer with a six-nucleotide index and 2 uL Taq Gold, in a total volume of 100 uL. qPCR was performed on the libraries to assess the quantity of DNA. The PCR conditions were followed according to manufacturer's directions, amplifying for 10-14 cycles, depending on the qPCR results. The PCR reaction was purified using the QIAQuick PCR Purification Kit, with elution in 30 uL EB Buffer and an incubation at 37 C for 10 minutes prior to the elution step. DNA concentration was assayed using a Qubit 2.0 Fluorimeter, following

manufacturer's instructions. If the DNA concentration was less than 20 ng/uL, a second PCR amplification was performed using Phusion. The mix included 20 uL template DNA, 2 uL each of primers IS5 and IS6, 50 uL Phusion Master Mix, and 26 uL H2O. The PCR program followed manufacturer's instructions but for 6-10 cycles, and was purified with a QIAQuick PCR Purification Kit as described above. The capture procedure was performed following manual version 2.3.1 at the University of Copenhagen, with a 65 C incubation for 18 hours. The heat elution step was skipped, and the capture was amplified using KAPA Hi-Fi polymerase, following manufacturer's instructions for 16 cycles. The PCR reaction was cleaned using a QiaQuick PCR Purification Kit, eluting 30 uL of EB Buffer after a 15-minute incubation at 37 C. The capture was visually examined and quantitated using an Agilent 3300 Bioanalyzer. If the DNA concentration was lower than 20 ng/uL, the capture was reamplified using the KAPA polymerase.

Samples were pooled 8-10 individuals to a sequencing lane, and were sequenced on an Illumina HiSeq 2500. The samples captured at Copenhagen were sequenced at the Danish National DNA Sequencing Center, and the samples captured at the University of Illinois were sequenced at the Roy J. Carver Biotechnology Center at the University of Illinois. The sequences were then run through a computational pipeline developed by the Malhi lab, which uses AdapterRemoval to trim adapters off of the reads (Lindgreen, 2012), bowtie2 to align the reads to the dog mitogenome reference (Genbank Accession NC002008) (Langmead and Salzberg, 2012), SAMtools to remove duplicates (Li et al., 2009), SNVer to call variants (Wei et al., 2011), and MapDamage to analyze the sequence reads for damage patterns consistent with ancient DNA (Jónsson et al., 2013). All deduplicated sequence files were examined by eye in Geneious version 8.1 (http://www.geneious.com, Kearse et al., 2012), to confirm all variant calls and look for contamination or nuclear inserts. SNP calls were determined by comparing the results generated by SNVer to the alignment of reads in Geneious. For a SNP to be confirmed, it had to be present in at least 67% of the reads. If a SNP was found in reads that seemed to be duplicates of one another (but that had differing levels of damage on the ends so they were not considered "identical" when

duplicates were removed), that was considered a single instance of the SNP. If a SNP was found in 33% or less of the reads, the SNP was not counted. If the called SNP was present in 33-67% of the reads, or if it was found in a region of 1X coverage, it was left as an ambiguous base. All SNPs were treated as independent.

If a read differed by more than three base pairs from the reference sequence, that read was compared to the database of published DNA sequences maintained by the National Center for Biotechnology Information to confirm its identity using Web BLAST (Altschul et al., 1990, https://blast.ncbi.nlm.nih.gov/Blast.cgi). If the read matched most closely to human DNA, dog nuclear mitochondrial insertions, or the DNA of another species, it was removed. This screening also identified the mitogenome from the Anker site as belonging to a coyote, and it was also removed from further analysis. Five additional samples were removed from the analysis due to the low number of dog mitochondrial reads recovered – samples with less than 15,000 bp coverage and 5X sequencing depth were excluded. In all cases, these six samples had a high rate of duplication and low coverage – on average, 19% of the mitochondrial genome had no coverage, and 37% of the mitochondrial genome had only 1X coverage. The samples retained for analysis had an average of 56X coverage of the mitochondrial genome, ranging from 7X to 122X. The sequencing results for each sample is provided in Table 3.2, and the full list of variants identified in all samples is provided in Table 3.5.

The ancient dog mitogenomes were combined with a dataset of 28 modern dogs (Ishiguro and Nakano, 1996; Björnerfeldt et al., 2006; Pang et al., 2009; Webb and Allard, 2009; Verscheure et al., 2014; Duleba et al., 2015). Modern dogs are divided into six major mitochondrial haplogroups, many of which have subhaplogroups (haplogroup A, A1, A1a, etc.). One modern dog from each haplogroup and subhaplogroups down to the third level (e. g. A1a, A1b, A4a) was selected as a representative for their subhaplogroups (Table 3.3). A second dataset of modern dogs was assembled with representatives for every subhaplogroup in haplogroup A (which is by far the most common haplogroup in dogs worldwide),

as a more complete dataset of 103 individuals (Table 3.4). The more complete dataset did not affect the results or their interpretation. In both cases, where possible, the haplotype in the dataset derived from an American modern dog, to maximize the likelihood that shared haplotypes between ancient and modern dogs could be identified. In addition to the modern dogs, ancient and modern wolves (Thalmann et al., 2013; Loog et al., 2017) were included in the analysis, as well as coyotes (Björnerfeldt et al., 2006) to serve as outgroups. All sequences were aligned using Mafft version 7 (Katoh et al., 2002; Katoh and Standley, 2013) on the CIPRES Science Gateway (Miller et al., 2010). The dog mitochondrial genome has a 10-bp repeat region that spans nucleotide position 16131 to 16499 (Kim et al., 1998). Because this region is hard to sequence and was often absent from the sequences used for comparison, and because it is hypervariable, and therefore likely saturated with substitutions, this region was removed from the alignment prior to further analyses.

A maximum likelihood tree was built from this alignment using RAxML version 8 (Stamatakis, 2014) on the CIPRES Science Gateway (Miller et al., 2010). A rapid bootstrap analysis was used, with 500 bootstraps and a GTR-gamma model of nucleotide evolution. All trees were visualized in Figtree v. 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). Multiple median-joining haplotype networks (Bandelt et al., 1999) were built using PopART (http://popart.otago.ac.nz). A network was constructed comparing ancient and modern dogs, to assess the relationship among haplotypes and the possibility of ancient dog lineages being shared with or similar to haplotypes from present day dogs. A second network consists solely of ancient dogs, to show how dogs from different time periods and archaeological sites were related. Population diversity measures, such as nucleotide diversity and Θ_S, were calculated using Arlequin (Excoffier and Lischer, 2010).

Demographic Modeling

To model the demographic history of dogs in the Americas, all ancient and modern wolves, modern dogs, and the 11 Siberian individuals were excluded. PartitionFinder was used to determine if the

data needed to be partitioned into different models for different regions of the mitogenome (Lanfear et al., 2012). A greedy search scheme was performed using Bayesian Information Criteria, and each gene was considered separately (Lanfear et al., 2012). A total of three partitions were identified in the data, and can be found in Table 3.7. The partitioned alignment file was then analyzed using BEAST 2, to construct an extended Bayesian skyline plot, which uses a Markov chain Monte Carlo method to identify coalescence times for a population, and uses the timing of those coalescences to estimate demographic changes over time (Drummond et al., 2005; Bouckaert et al., 2014). Unless mentioned, the default parameters were used. The tip dates were estimated as the midpoint of the age estimate for each individual. An estimated lognormal clock was used with a given range of $1x10^{-5}$ and $1x10^{-10}$, with a starting value of $1x10^{-8}$. The population mean prior was lognormal, with a mean of 5 and a standard deviation of 1.5. The Markov chains were run for 100 million generations, and the first 10% was discarded as burn-in. The Bayesian Skyline Plot was graphed in R, and the resulting tree was visualized and formatted in Figtree v. 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

To support the results of the demographic model, a simulation showing stable population size over time was constructed using BEAST version 1.8.4 (Drummond et al., 2012). The simulated dataset was constructed to have the same number of taxa, and the same length, nucleotide frequencies, tip dates, and number of variant sites, to keep the conditions between the two datasets as similar as possible. The simulated dataset was set to reflect a stable population size of 20,000 individuals starting at 20,000 ybp, with a much lower population of 1,000 individuals prior to that. The simulated data was not partitioned, and the demographic model parameters from Subset 1 in the original analysis were used in this case. The parameters used in the analysis were the same, although the Markov chains were run for 250 million generations in this case, with the first 10% discarded as burn-in. The skyline plot was graphed using Tracer version 1.6 (http://tree.bio.ed.ac.uk/software/tracer/).

Results

All ancient American dogs are members of the modern dog haplogroup A, as shown in Figure 3.3. However, the ancient dogs do not belong to any specific subhaplogroups within A, and there is no haplotype sharing between ancient and modern dogs (Figure 3.5). Some individual haplotypes, including those part of the Southeast and Midwest clade, are basal to haplogroup A, while other haplotypes from Siberia, Argentina, and the Yucatan are more closely related to modern subhaplogroups, but differ by more than fifteen substitutions from a modern dog subhaplogroup. Other ancient haplotypes that are more than thirty base pairs from a modern dog subhaplogroup, including some Western dog haplotypes, form their own mitochondrial clade. The ancient Siberian dog haplotypes are also part of modern dog haplogroup A, but form their own clades, separate from the ancient American dogs (Figure 3.3). The wolves that are most closely related to the ancient American dogs are from Europe and Siberia (Figure 3.4).

The demographic model of the ancient American dogs suggested that the population of dogs in the Americas has coalescence dates of 11,400 to 17,600 ybp (Figure 3.9). After migrating to the Americas, the dog population gradually increased until 8000 ybp, and then plateaued. The population size then showed a decline starting at 2000 ybp (Figure 3.8). Conversely, the simulated dataset with stable population size shows no change in population over time (Figure 3.10). The demographic model also reveals two major clades of dogs in the Americas, each with deep divergence times (Figure 3.9). One clade contains some Alabama individuals and nearly all of the dogs from the Midwest sites of Scioto Cavern and Janey B. Goode. The other clade is more geographically widespread, and includes some individuals from the Southeast, the Midwest, and the dogs from the West coast of the United States, the Yucatan, and Argentina. The haplotypes that are part of the Southeast/Midwest United States clade are genetically very similar, and differ from each other by less than ten base pairs (Figure 3.7). The clade that is more geographically widespread is also more genetically diverse, including haplotypes from Florida, Argentina,

and Apple Creek, Illinois, that differ from all other haplotypes sampled by at least eight substitutions. As shown in Figure 3.3, they still lie within dog haplogroup A, suggesting that these haplotypes are not likely to be the result of admixture with wolves. These haplotypes may be American founding haplotypes, as they are too divergent from the other haplotypes in the sample to have arisen due to random substitution over time.

The different dog populations vary in genetic diversity levels, as shown in Table 3.6. The variation in genetic diversity does not seem to follow a clear temporal or geographic trend, and there is no correlation between sample size and genetic diversity (R^2=.0826, p = 0.421) (Figure 3.6). Additionally, there are no clear patterns for how genetic diversity changes across time or space. However, in the Midwestern United States, there is a clear pattern of increasing genetic diversity in dogs over time (Figure 3.11). Haplotypes from the older populations (Koster, Modoc, and Scioto) cluster together, with six or fewer substitutions between them, as shown in Figure 3.12. However, haplotypes from the more recent populations (Janey B. Goode, Angel Mounds, and Apple Creek) are much more divergent, and differ from one another by as many as 20 substitutions.

Discussion

Ancient American dogs were not found to share haplotypes with living dogs. This is in contrast to previous studies, which have suggested that some village dogs and isolated breeds share mitochondrial DNA with ancient dogs (Brown et al., 2013; van Asch et al., 2013). As we showed in Figure 3.1, it is possible that multiple mitogenome haplotypes have identical HVR haplotypes. It is also possible that shared ancestry could be identified if more mitogenomes are sequenced for isolated populations. The modern dog haplogroup that is most similar to ancient dog mitogenome haplotypes is A4, which differs from ancient dog haplotypes from the Yucatan and Argentina by roughly 20-25 substitutions. Haplogroup A4 contains only two haplotypes (A4 and A4a) and has only been identified in six published dog mitogenomes (Table 3.8). Five of these dogs derive from Asia, from historically Asian breeds (such as the Pekingese and

Japanese Spitz), or both. The sixth dog is a Chihuahua from the United States, which is interesting because chihuahuas have been suggested to have ancient American origins (van Asch et al., 2013; Parker et al., 2017). This suggests that the haplotypes from modern haplogroup A4 are similar to the dog haplotypes from Central and South America, and ancient dogs with a similar haplotype moved with humans from Beringia into the Americas.

The lack of shared haplotypes suggests that very little ancient American dog ancestry, if any, remains in living dog populations. Domestic dogs experienced large population bottlenecks, both during domestication and breed formation (Lindblad-Toh et al., 2005; Karlsson et al., 2007; Larson et al., 2012; Freedman et al., 2014). The ancient dogs only experienced one of these bottlenecks, so they would have higher genetic diversity. Most living dogs that have been sampled for mitochondrial DNA are breed dogs, and have lower genetic diversity than village dogs, which have not experienced the same population bottlenecks (Boyko et al., 2009; Shannon et al., 2015). It should be noted that mitogenomes are maternally-inherited, and therefore represent just a fraction of each individual's ancestry (Richard E Giles et al., 1980). Nuclear genomic regions should be sequenced to test this hypothesis.

Although previous work has suggested that dogs may have arrived in the Americas as recently as 10,000 ybp (Witt et al., 2015), analysis of complete mitogenomes indicates that the coalescence time for dog populations in the Americas is consistent with the timing of human arrival in the Americas, which has most recently been estimated at around 16,000 ybp (Achilli et al., 2008; Kemp and Schurr, 2010; Meltzer and Holliday, 2010; Llamas et al., 2016). The demographic history of dogs as reflected in the extended Bayesian skyline plot shows a long period of population stability, followed by a decline that started around 2000 ybp. This decline becomes sharper around 1000 ybp (Figure 3.8). However, the confidence intervals are broad, and so the decline may not be as dramatic as it appears. However, when compared with a simulated stable population, there are clear differences, so it is likely that the shifts in population size are not just an artifact of the data (Figure 3.10).There is a similar decline in the archaeological record, which

shows a decrease in dog burials over time in multiple regions, including in the Mesa Verde region of the Southwest after 900 ybp (Emslie, 1978), the Midwest after 1000 ybp (Lapham, 2010), and the Southeast after 3000 ybp (Morey, 2006). In some cases, this shift also coincided with changes in dog usage – perhaps as new cultures developed and people adapted to new lifestyles, the need for dogs decreased and their importance in the community decreased as well. If humans were no longer actively provisioning dogs, that may have limited dog population size in a region, decreasing diversity. It's also possible that humans were still using dogs, but stopped burying them, so a decrease in the presence of dog burials over time does not necessarily represent a decrease in the dog population in the Americas. Alternatively, if humans were breeding dogs more intensively, then increased population structure and higher rates of loss to genetic drift could also have contributed to the decrease in effective population size.

While the Siberian dog population is thought to derive from the same ancestral dog population as the American dog population, the Siberian dog haplotypes differed from the American dog haplotypes by at least six substitutions, suggesting that there was a period of isolation after these populations split. This mirrored a similar pattern in Native American mitochondrial haplotypes, which are similar to but divergent from Siberian mitochondrial haplotypes (Torroni et al., 1993; Forster et al., 1996; Schurr and Sherry, 2004; Tamm et al., 2007; Malyarchuk et al., 2011). A hypothesis for why the haplotypes between Siberians and Native Americans are distinct from one another, known as the Beringian Incubation Model, suggested that human populations spent a prolonged period of time on Beringia before moving into the Americas (Tamm et al., 2007; Kitchen et al., 2008; Mulligan et al., 2008; Llamas et al., 2016). Additionally, the ancient American dogs are most closely related to ancient wolves from Siberia and Switzerland, rather than American wolves (Figure 3.4). This pattern also supported the hypothesis that dogs migrated with humans to the Americas and were part of the Beringian Incubation, rather than being domesticated from North American wolves (Leonard et al., 2002) after Native Americans migrated to the Americas.

Interestingly, some of the patterns found in the dog populations mirrored what we know of Native

American population history. For example, two major clades have been identified in ancient Native American populations, and are known as the Northern and Southern clades (Achilli et al., 2013; Rasmussen et al., 2014; Verdu et al., 2014). One clade was found in individuals from both North and South America, while the other clade was limited to individuals from North America, especially in the Arctic (Rasmussen et al., 2014). Dogs in the Americas also fit into two mitochondrial clades, each with a coalescence time of 14,000-10,000 ybp. In dogs, one clade consists of Southeast and Midwestern dogs, and the other has a much broader geographic distribution. The "Northern" clade of Southeast and Midwest dogs is star-like, and we calculated Tajima's D for this clade using Arlequin, with 1000 simulations (Excoffier and Lischer, 2010). The clade had a Tajima's D of -2.035 (p=.006), suggesting that the Northern clade was expanding in population size over time. In almost all cases, all haplotypes identified from a single archaeological site are found within a single clade. The exceptions are Angel Mounds, with two individuals in the widespread clade and two in the Southeast/Midwest clade, and Janey B. Goode, with three individuals in the widespread clade and fifteen individuals in the Southeast/Midwest clade. These populations (which have higher genetic diversity in Table 3.6) are both from the Midwest, during the Terminal Late Woodland and Mississippian periods, when there was an increased amount of trade in the region (Brown et al., 1990; Kelly, 1990; Pauketat, 2004; Smith, 2007). It is possible that these dogs derive from other regions of the Americas, and were then traded to the Midwest. Dogs from the Channel Islands cluster both with dogs from British Columbia and from the Southwest, in two separate subclades, suggesting that the diversity of the dog population on the Channel Islands may also be a result of human trade interactions (Rick et al., 2005, 2008; Gill and Erlandson, 2014).

At a regional scale, there are two examples of dog demography mirroring human demography: on the Pacific coast, and in the eastern United States. Haplotypes from dog populations along the Pacific coast cluster together (Figure 3.3). One cluster consists of individuals from Prince Rupert Harbour and British Columbia and one dog from the Channel Islands (Figure 3.7). This is consistent with a human

migration route along the Pacific coast (Eshleman et al., 2004; Fix, 2005; Perego et al., 2009; Reich et al., 2012; Battaglia et al., 2013), with limited admixture with inland populations (Eshleman et al., 2004; Verdu et al., 2014). However, most of the Channel Island dogs cluster instead with dogs from the Southwest, suggesting that further south, intermixing between dog populations from the coast and inland is more common. The people living on the Channel Islands had widespread exchange networks with the mainland, and in some cases trade items found on the Channel Islands originated from places as distant as modern-day Nevada or Oregon (Rick et al., 2005; Perry, 2012). It has also been suggested that dogs were initially introduced to the Channel Islands via trade (Rick et al., 2008), so it is possible that additional dog haplotypes were introduced to the islands through trade as well. In the Midwest and Southeast, individuals from archaic sites in Alabama (including Perry, Flint River and Little Bear) cluster with multiple populations in the Midwest, including Janey B. Goode, Koster and Scioto Caverns. The Southeast and the Midwest Native American populations were often part of the same cultural interaction sphere, including during the Hopewell (Struever and Houart, 1972; Brose and Greber, 1979; Seeman, 1979; Charles and Buikstra, 2006) and Mississippian (Kelly, 1990; Pauketat and Emerson, 1997; Smith, 2007; Pauketat and Alt, 2015) periods. Similarities between the dog populations from these regions are consistent with this shared cultural history. Interestingly, the Mississippian individual from Florida was highly divergent from this clade, and was not close to any other individual, suggesting that the dog population in the Southeast may have been replaced with dogs from another region after the Archaic period.

Connections can also be made between dog demographic changes and human cultural changes. The Midwest is the geographic region with the largest sample size (N=35) in this study, and the dogs from this region range in age from 9000 ybp to 600 ybp. Therefore, we were able to examine the demographic history of this region in more detail. In the Midwest, dog populations increase in genetic diversity over time, especially in the period between 1200 and 800 ybp (Figure 3.11), in contrast to what we see overall in the Americas through the Bayesian skyline plot analysis. Interestingly, this increase in dog population

diversity also corresponds to the Woodland-Mississippian transition. The Late Woodland archaeological period ended and the Mississippian period began roughly 1000 ybp, and was marked by the founding of a large cultural empire that was centered around modern-day St. Louis, but extended throughout much of the Midwestern and Southeastern United States (Emerson, 1997; Pauketat and Emerson, 1997; Emerson and Lewis, 1999; Pauketat, 2004). The shift from Late Woodland to Mississippian periods the Mississippi river valley in Southern Illinois resulted in a number of lifestyle changes in human population, including an increase in population density and village size, a shift from hunting and gathering with small-scale horticulture to maize agriculture, and the rise of Cahokia, a large mound city with far-reaching cultural influences (Emerson, 1997; Pauketat and Emerson, 1997; Pauketat, 2004; Alt, 2010).

The increase in dog diversity in the Midwest between 1200 and 800 ybp could have two causes. First, it is possible that there were many migrants to the Mississippian empire, and these migrants could have brought dogs with haplotypes not previously found in the region. Ancient DNA studies have indicated some continuity between Late Woodland and Mississippian populations (Raff et al., 2011; Reynolds et al., 2015), but analysis of strontium stable isotopes in human remains has shown that a number of Mississippian residents were not born near Cahokia (Slater et al., 2014). Second, the Mississippians had long-distance trade interactions, both within and outside of the empire (Brown et al., 1990; Pauketat and Emerson, 1997; Pauketat and Alt, 2015). Dogs may have been transported long distances through trade, which would also introduce new haplotypes to the existing dog population. In either case, the newly introduced dogs could have interbred with the existing population, resulting in higher levels of genetic diversity or introducing new alleles to the population. Interestingly, this increase in dog diversity also coincides with a shift in dog usage in the Midwest. Dogs in the Woodland period were often deliberately buried (Parmalee et al., 1972; Cantwell, 1980; Galloy, 2010), and some of them were likely used for hauling supplies on travois (Galloy, 2010). During the Mississippian period, however, dogs were no longer buried but instead were recovered in fragments from middens and trash piles (Schwartz, 2000). They were

primarily used as a food source, for feasts, during this time period (Schwartz, 2000; Galloy, 2010). Perhaps this shift in dogs' roles in human society also had an impact on its population size and structure.

Conclusions

This study is one of the first comprehensive analyses of ancient dog mitogenomes in the Americas, and vastly expands the number of published mitogenomes available for study, from 3 to 71. In this study, we tested two hypotheses that were supported by HVR data, and rejected the hypothesis that ancient dog mitogenome haplotypes can be found in living dogs, and supported the hypothesis that dogs in the Americas were brought over with humans from Eurasia. We found that contrary to previous analysis, dogs arrived to the Americas during the initial human peopling. We identified that the dog population in the Americas had two primary clades with a deep divergence time of 13,000 to 17,000 ybp, similar to the Northern and Southern clade identified in Native American genomes, and that dog populations in the eastern United States and the Pacific coast show similar genetic relationships to human populations in the area. Finally, we found that with enough sampling of dogs across time and space, it is possible to identify changes in the dog population that coincide with human cultural changes, such as the Mississippian transition. By examining dog demographic history in the Americas, we may better understand human history in the Americas, as well as how dogs and humans moved and interacted in the Americas.

Figures and Tables

Figures

Figure 3.1: Median joining networks of mitochondrial DNA of ancient dogs in the Americas. The same individuals are on both networks. 3.1A shows a network of complete mitogenome sequences, while 3.1B shows a network of HVR sequences. This illustrates that by sequencing complete mitochondrial genomes, one can see greater haplotype differentiation.

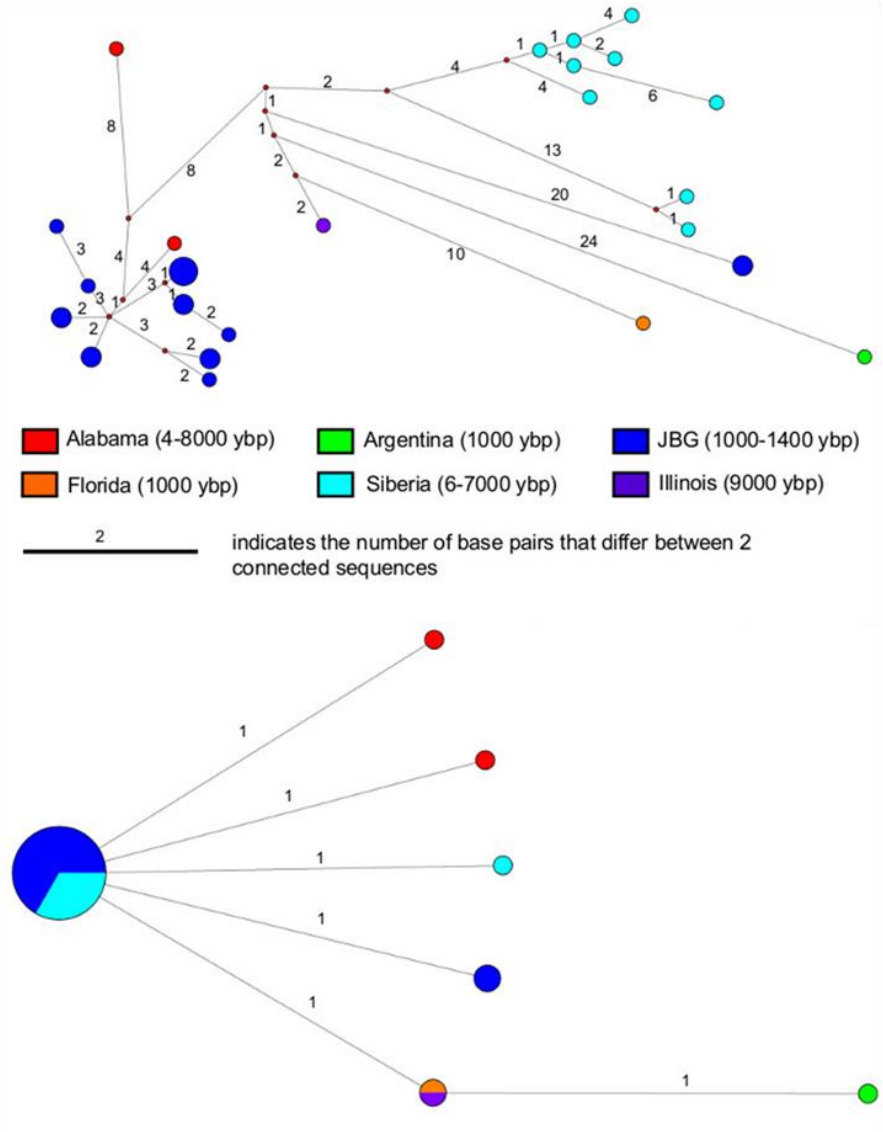


Figure 3.2: A map of the samples used in this study. Each circle represents an archaeological site, or multiple archaeological sites from the same area and time period, and they are color-coded by the age of the samples. The numbers indicate the site number, and a brief guide to each site's age, location, and number of samples represented is included here. 1. Aachim Lighthouse, Siberia. 2 samples, 1750 ybp. 2. Zhokhov, Siberia. 8 samples, 8000 ybp. 3. Prince Rupert Harbour, British Columbia. 5 samples, 1500 ybp. 4. Channel Islands, California. 8 samples, 2000-5000 ybp. 5. Grass Mesa, Colorado. 1 sample, 1100-1400 ybp. 6. McPhee Pueblo, Colorado. 1 sample, 1100-1300 ybp. 7. Yellow Jacket Pueblo, Colorado. 1 sample, 800-1000 ybp. 8. Simonsen Bison Kill, Iowa. 1 sample, 7200-7600 ybp. 9. Janey B. Goode, Illinois. 19 samples, 1000-1400 ybp. 10. Apple Creek, Illinois. 1 sample, 1000-2500 ybp. 11. Anker, Illinois. 1 sample, 1000-1400 ybp. 12. Scioto Caverns, Ohio. 7 samples, 2000 ybp. 13. Angel Mounds, Indiana. 4 samples, 1000 ybp. 14. Koster, Illinois. 3 samples, 9000 ybp (one sequence previously published). 15. Modoc, Missouri. 2 samples, 9000 ybp. 16. Cox, Alabama. 1 sample, 1500-3000 ybp. 17. Flint River, Alabama. 1 sample, 3000-7000 ybp. 18. Little Bear, Alabama. 1 sample, 3000-7000 ybp. 19. Perry, Alabama. 3 samples, 3000-7000 ybp. 20. Florida. 1 sample, 1000 ybp (sequence previously published). 21. Mayapan, Mexico. 4 samples, 1000 ybp. 22. Cerro Lutz, Argentina. 1 sample, 1000 ybp (sequence previously published).

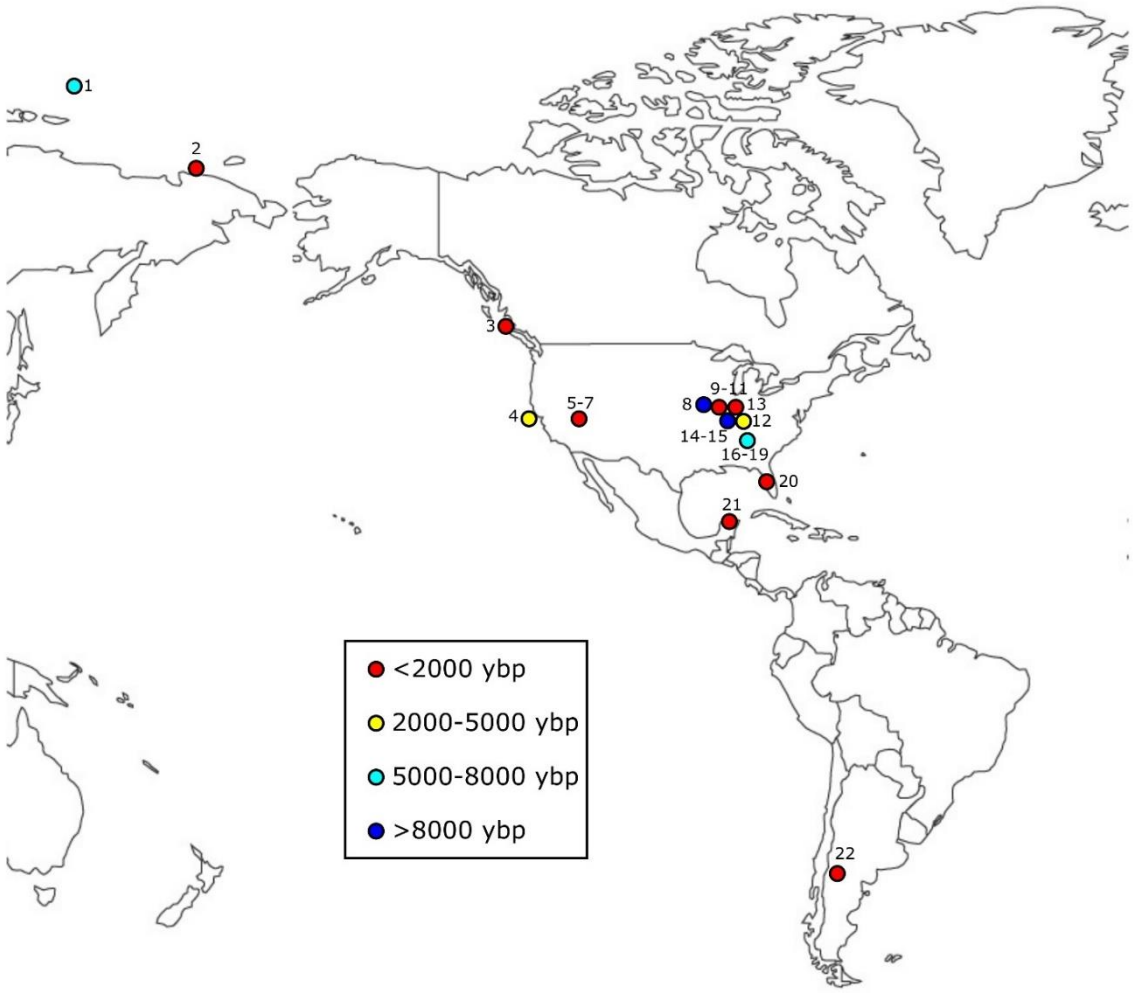


Figure 3.3: A Maximum Likelihood tree of ancient and modern dogs and wolves, with a coyote as an outgroup. Wolves and coyotes are in black, modern dogs are in gray, and ancient dogs are color-coded by their geographic region. “Ancient Siberian Dogs” includes dogs from the Zhokhov and Aachim Lighthouse sites, “Ancient Western North American Dogs” includes dogs from the Prince Rupert Harbour, Channel Islands, Grass Mesa, McPhee Pueblo, and Yellow Jacket Pueblo Sites”, “Ancient Central North American Dogs” includes dogs from the Janey B. Goode, Scioto, Angel Mounds, Apple Creek, Modoc, and Koster sites, “Ancient Eastern North American Dogs” includes dogs from Florida and the Perry, Cox, Little Bear, and Flint River sites, and “Ancient Central and South American Dogs” includes dogs from Mayapan and Cerro Lutz sites. Stars indicate nodes with a bootstrap value > 90%.

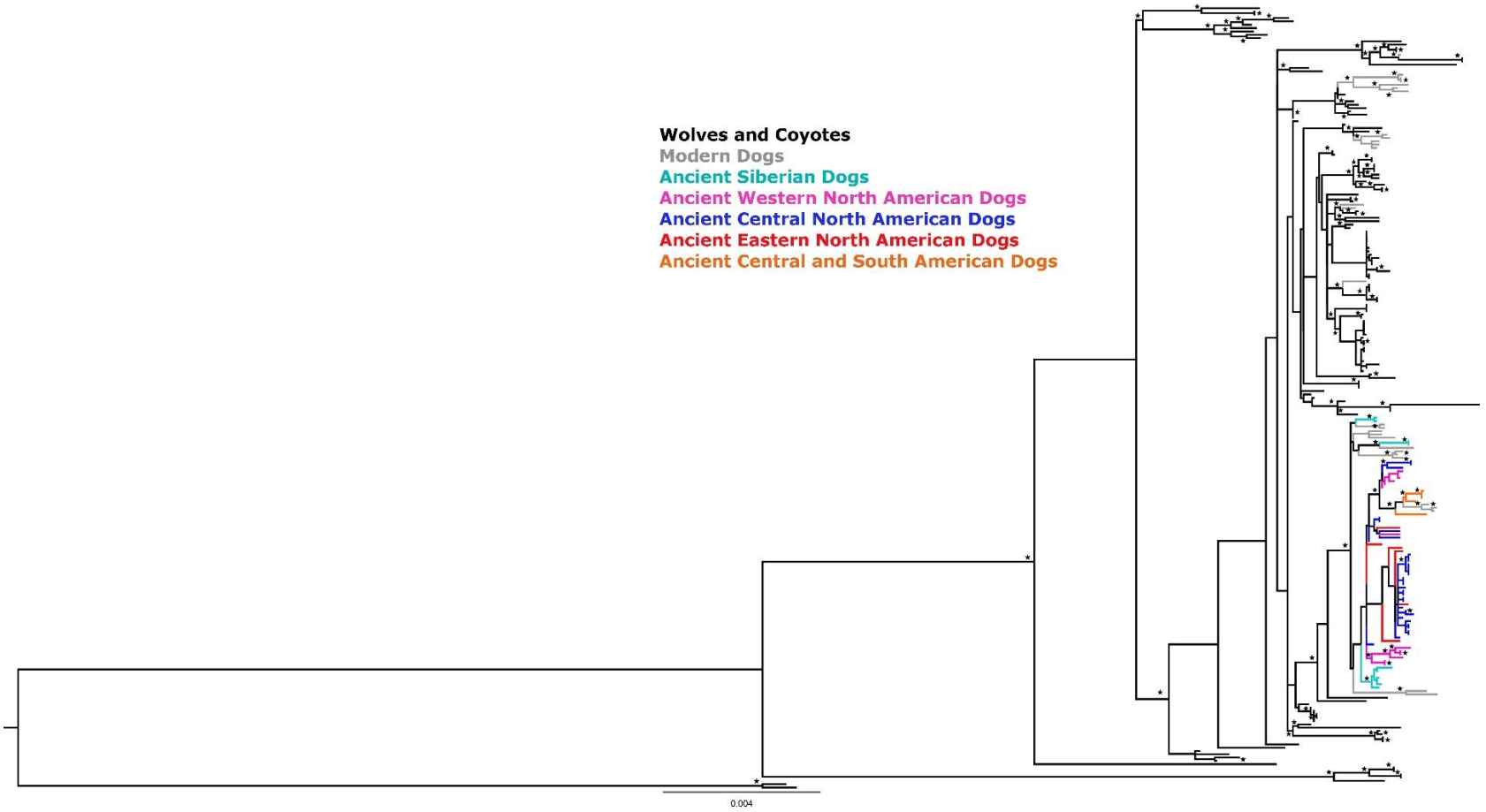


Figure 3.4: A Maximum Likelihood tree of ancient and modern dogs and wolves, with a coyote as an outgroup. Coyotes are in black, ancient dogs are in dark gray, modern dogs are in light gray, and wolves are color-coded by their geographic region. “European Wolves” includes wolves from Armenia, Belgium, Croatia, the Czech Republic, Denmark, Finland, Germany, Italy, Poland, Sweden, Switzerland, Spain, and Ukraine, “Russian and Siberian Wolves” includes wolves from Russia and Siberia, “Middle Eastern Wolves” includes wolves from Afghanistan, Iran, India, Israel, Oman, Saudi Arabia, Syria, and Turkey, “Asian Wolves” includes wolves from China, Japan, and Mongolia, and “American Wolves” includes wolves from Alaska, Canada, the mainland United States, and Mexico. Stars indicate nodes with a bootstrap value > 90%.

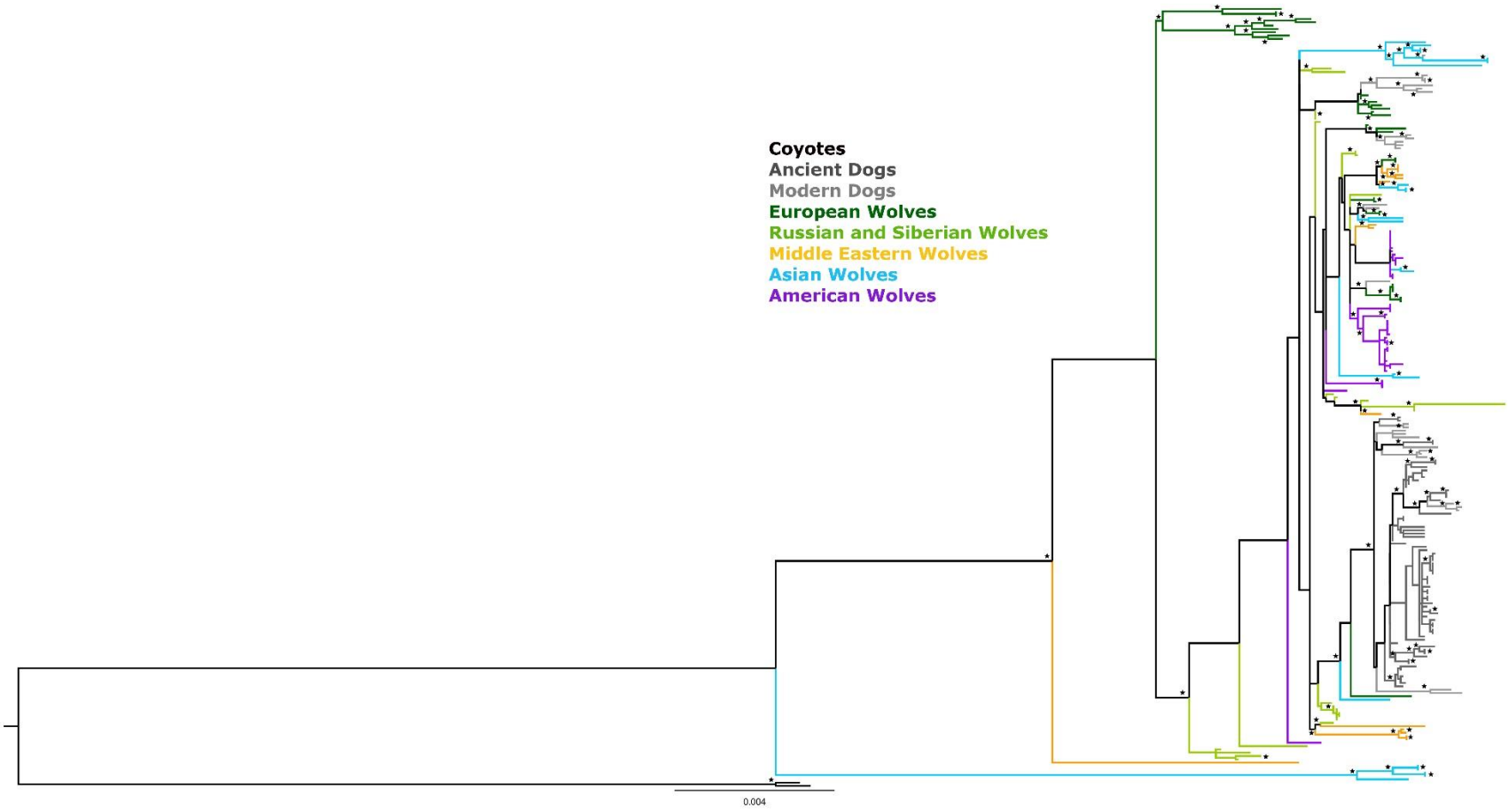


Figure 3.5: A median joining network of ancient and modern dogs. Each circle represents a unique haplotype, the size of the circle indicates the number of individuals with that sequence, and the number of tick marks indicate the number of base pairs that differ between sequences. The ancient dogs are color-coded by geographic region, while the modern dogs are gray. "Siberia" includes dogs from the Zhokhov and Aachim Lighthouse sites, "Western_Dog" includes dogs from the Prince Rupert Harbour, Channel Islands, Grass Mesa, McPhee Pueblo, and Yellow Jacket Pueblo Sites", "Central_Dog" includes dogs from the Janey B. Goode, Scioto, Angel Mounds, Apple Creek, Modoc, and Koster sites, "Eastern_Dog" includes dogs from Florida and the Perry, Cox, Little Bear, and Flint River sites, and "Mex_S_America_Dog" includes dogs from Mayapan and Cerro Lutz sites. The major modern dog haplogroups are labeled, as well as the sub-haplogroups of haplogroup A.

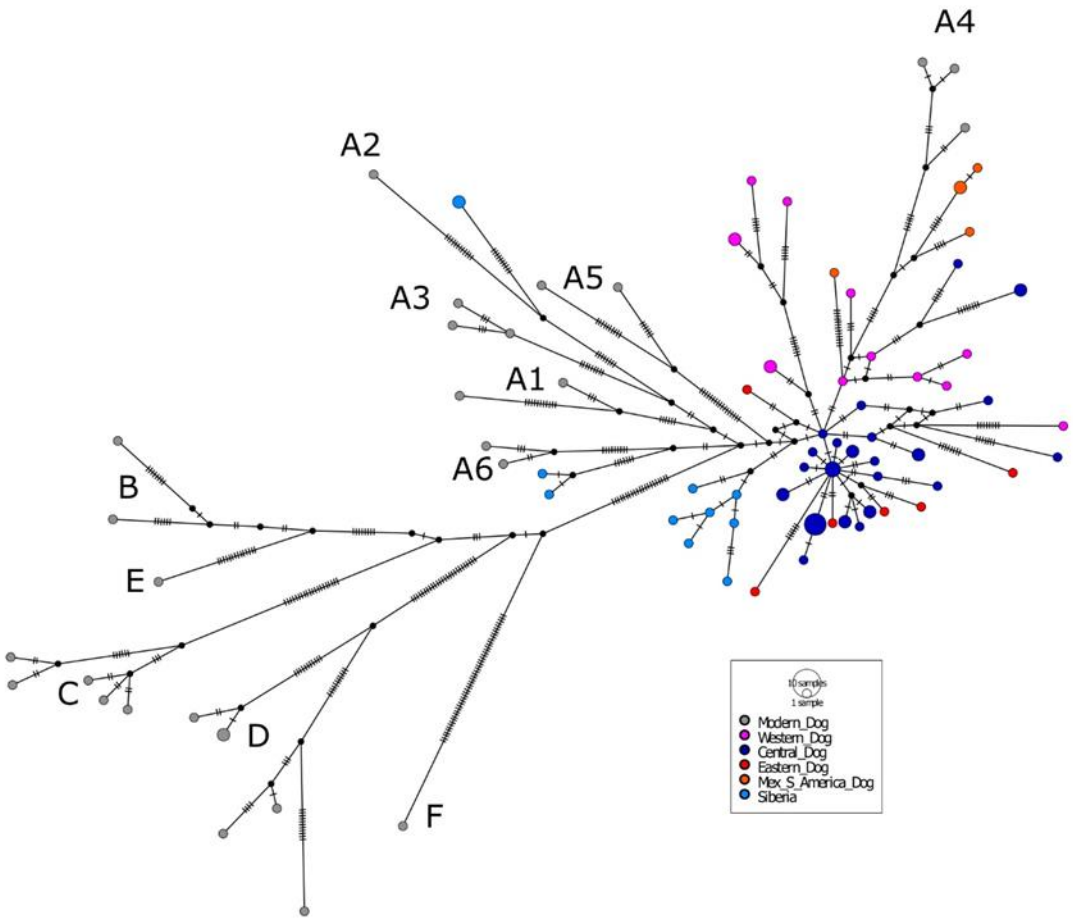


Figure 3.6: A plot showing the correlation between sample size and theta S, a measure of genetic diversity, for these populations. The R^2 value for the plot is shown above, demonstrating that there is a very weak relationship, which is confirmed by the p-value of 0.421.

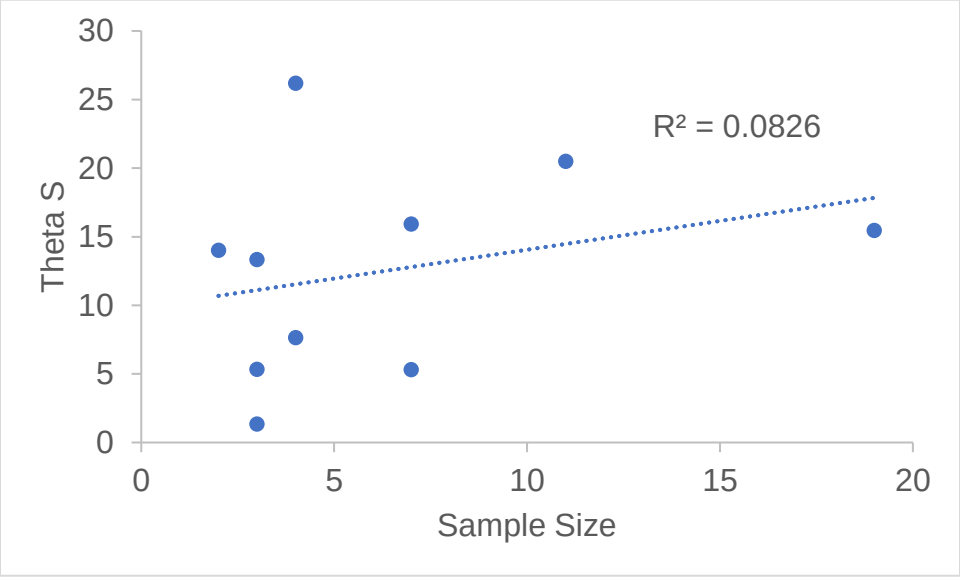


Figure 3.7: A median joining network of ancient dogs. Each circle represents a unique haplotype, the size of the circle indicates the number of individuals with that sequence, and the number of tick marks indicate the number of base pairs that differ between sequences. The ancient dogs are color-coded by archaeological site. Dogs from Grass Mesa, McPhee Pueblo and Yellow Jacket Pueblo are all considered "Southwest" in this figure.

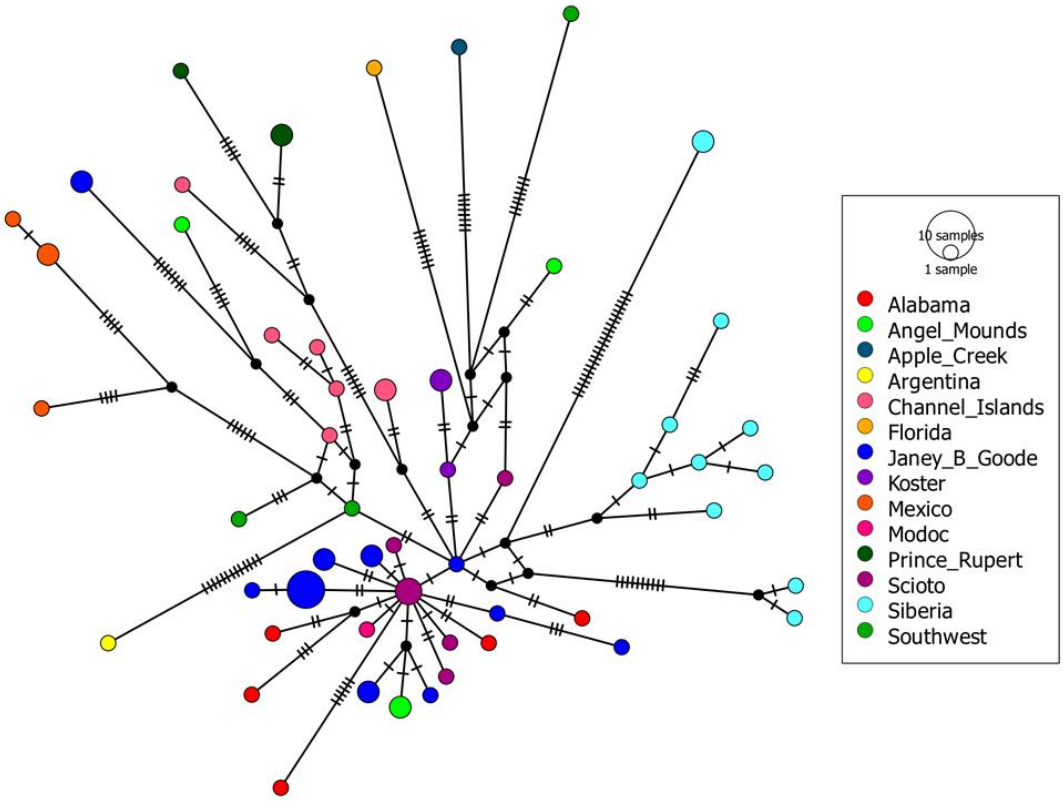


Figure 3.8: A Bayesian Skyline Plot of ancient dog mitogenomes in the Americas. The x axis is an indicator of time, with the present on the far left and going back in time on the right in years. The y axis is a measure of effective population size. The dotted line reflects the median value of the change in population size, while the gray shading indicates the 95% confidence interval.

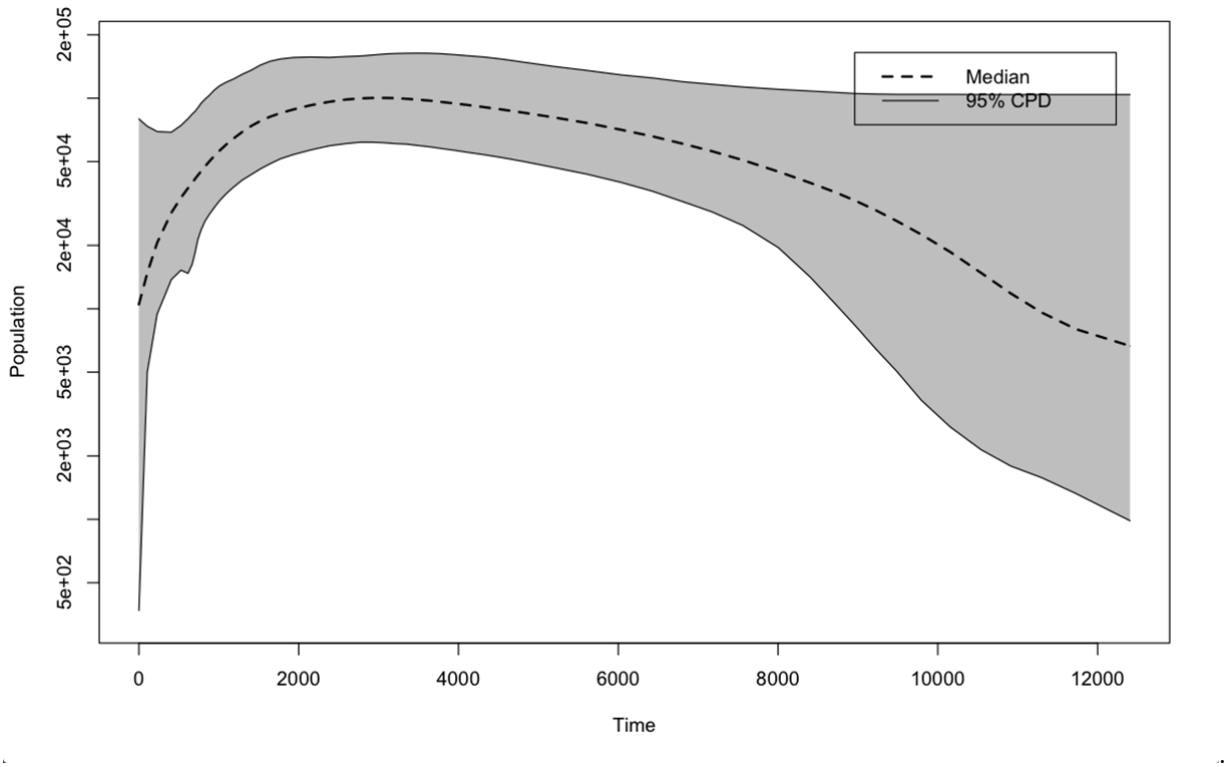


Figure 3.9: A Bayesian chronogram derived from the extended Bayesian skyline plot. The individuals are color-coded by geographic region, consistent with previous figures. The time scale is along the bottom of the graph, with the most recent time on the right and going backward in time as you move to the left. The light blue bars indicate the 95% confidence interval of the coalescence time.

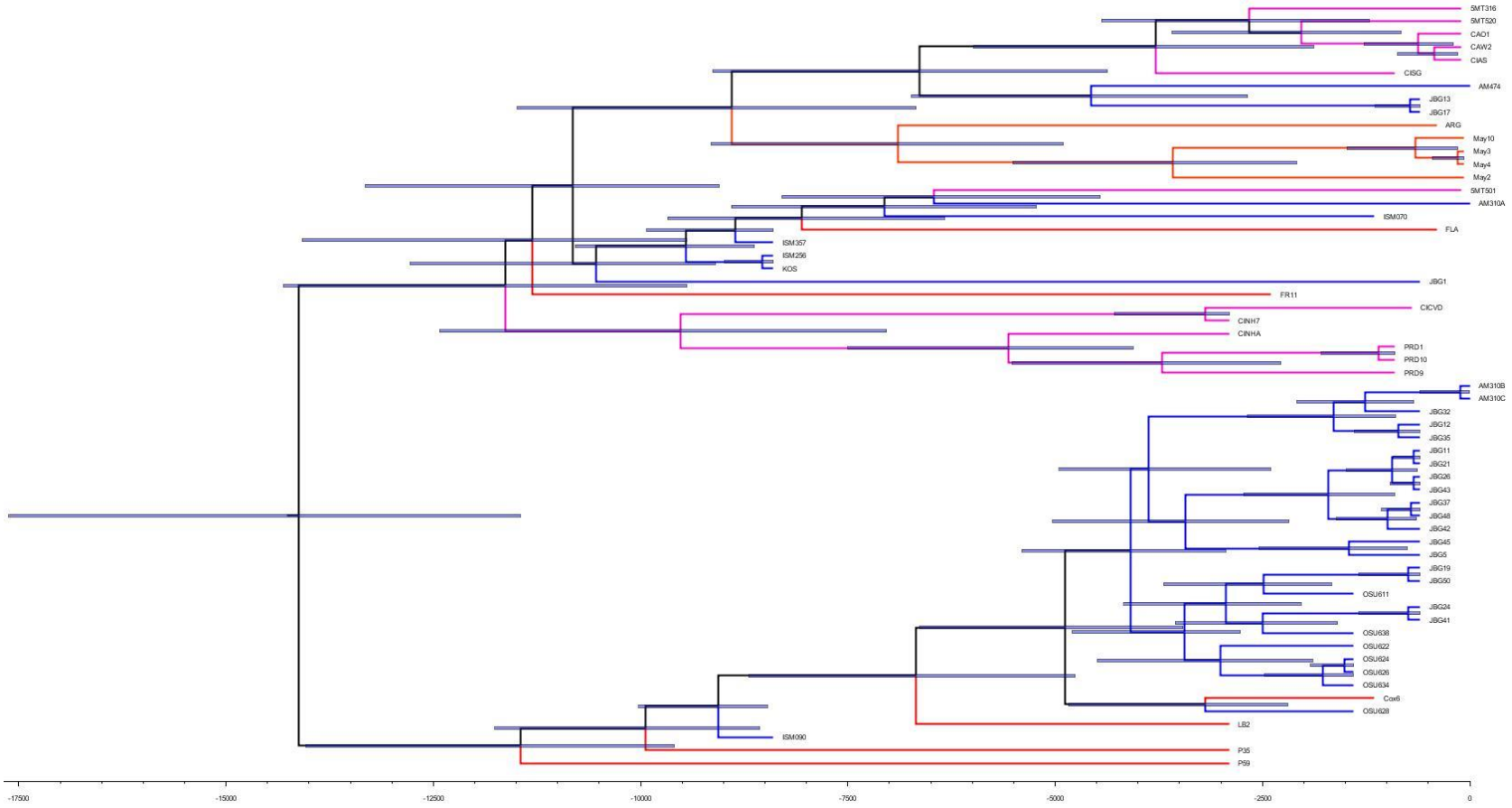


Figure 3.10: A Bayesian Skyline Plot of a simulated dataset (similar in parameters to the dog mitogenome sequences) of a stable population of 20,000 individuals for 20,000 years. The x axis is an indicator of time, with the present on the far left and going back in time on the right in years. The y axis is a measure of effective population size. The dotted line reflects the median value of the change in population size, while the blue shading indicates the 95% confidence interval.

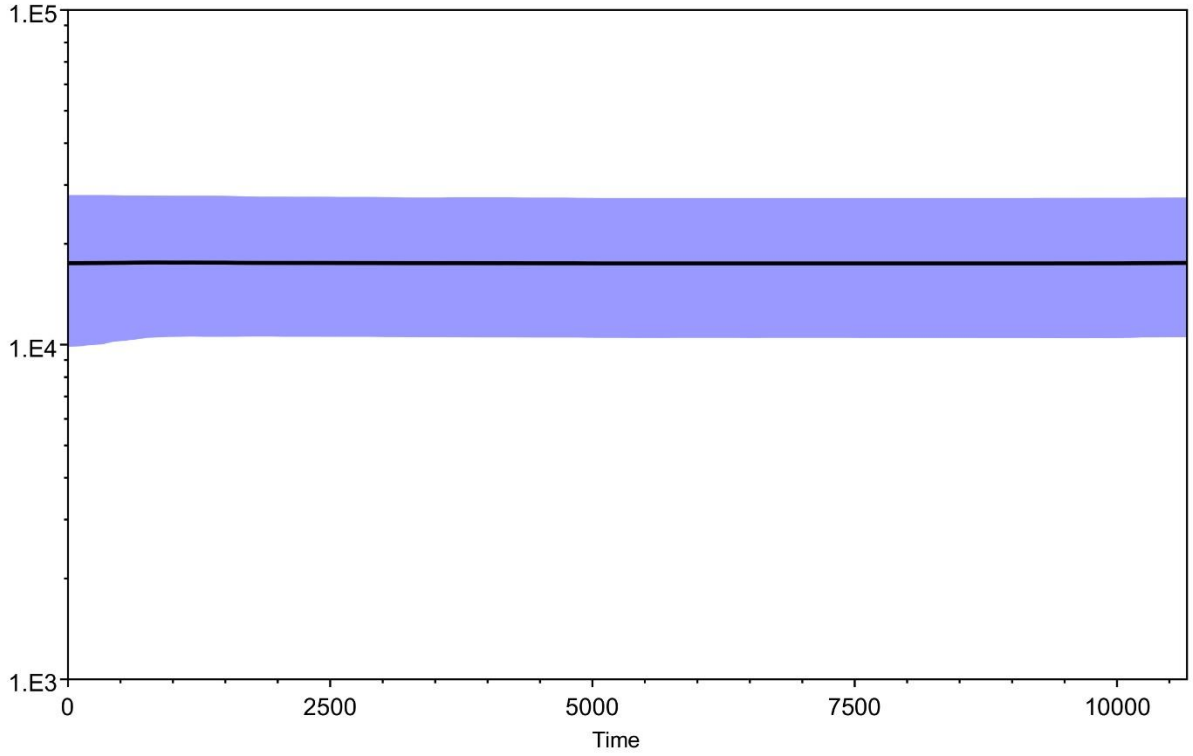


Figure 3.11: A plot of genetic diversity for all Midwest dog populations with N>1. This plot is of theta S and its standard deviation, with all values taken from Table 2.2. The sample size of each population is shown in parentheses next to the name, and the populations are in order from left to right of oldest to youngest population.

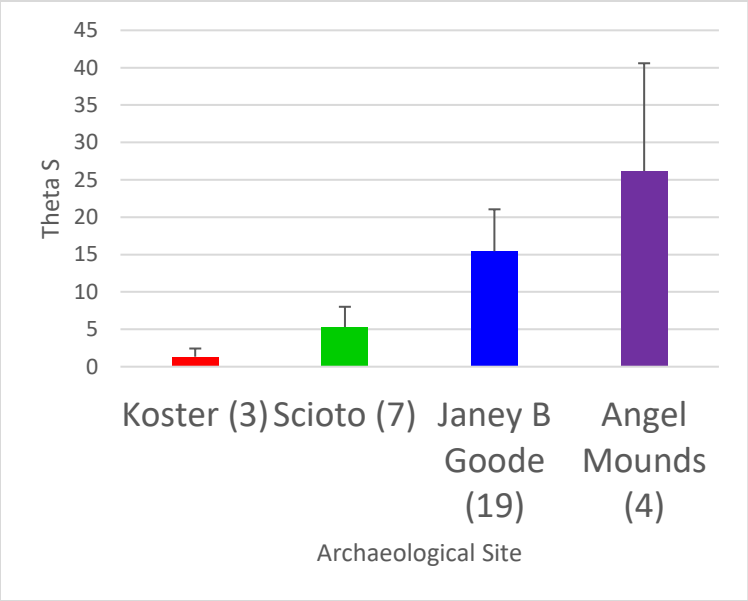
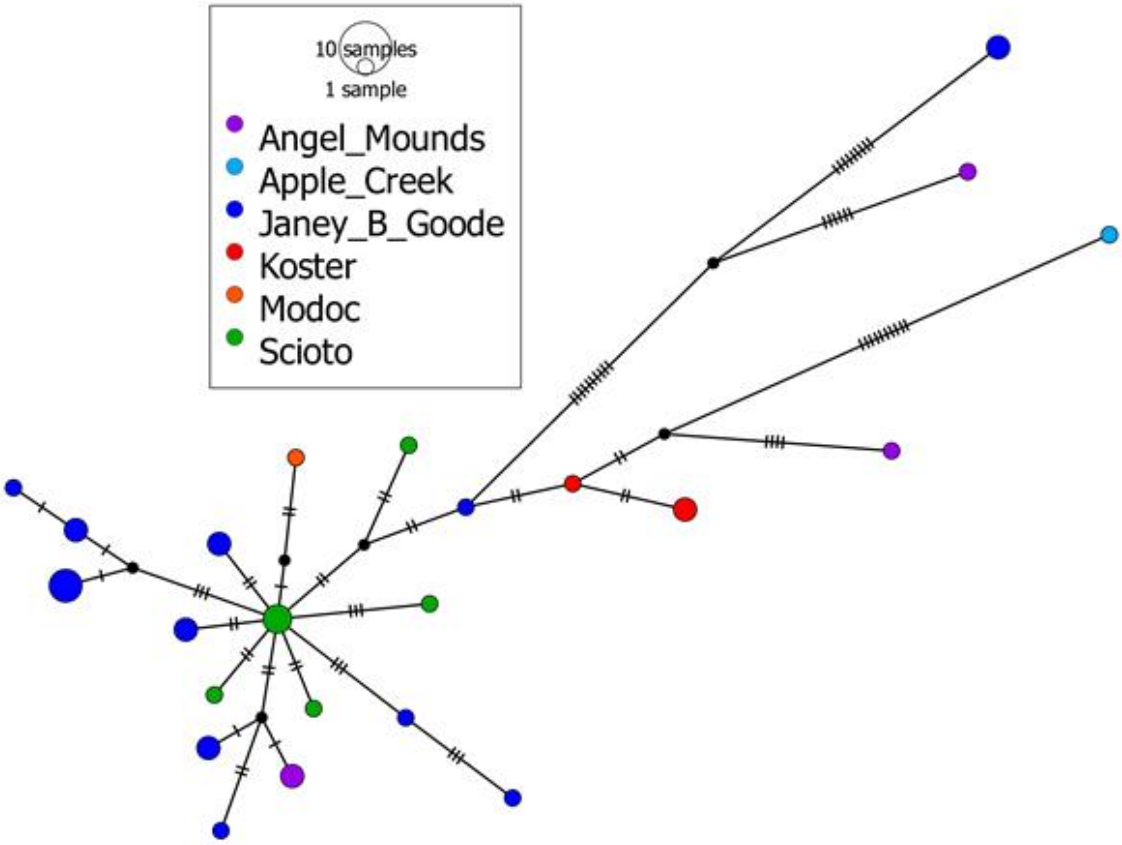


Figure 3.12: A median joining network of complete mitogenomes from the Midwest. Each circle represents a unique haplotype, the size of the circle indicates the number of individuals with that sequence, and the number of tick marks indicate the number of base pairs that differ between sequences. Each population is color-coded.



Tables

Table 3.1 A list of all ancient dog sequences analyzed in this study. The table lists the name of the archaeological site, the region the site is located in, the approximate age of the sample (determined by radiocarbon dating or provenience), the number of samples used from the site, and the reference in which it was first described. Sites marked with an '*' indicate that the sequence from this site was previously published. At the Koster site, 2 of the sequences are novel while one was already published.

Archaeological Site	Region	Age	Sample Size	Reference
Aachim Lighthouse	Eastern Siberia	1750 ybp	2	(Lee et al., 2015)
Angel Mounds	Indiana, Midwest	1000 ybp	4	Unknown
Anker	Illinois, Midwest	1000-1400 ybp	1	(Bluhm and Liss, 1961)
Apple Creek	Illinois, Midwest	1000-2500 ybp	1	(Parmalee et al., 1972)
Cerro Lutz*	Argentina	1000 ybp	1	(Acosta et al., 2011)
Cox	Alabama, East	1500-3000 ybp	1	(Moore, 1915)
Channel Islands	California, West	2000-5000 ybp	8	(Rick et al., 2008)
Flint River	Alabama, East	3000-7000 ybp	1	(Webb and DeJarnette, 1948a)
Florida*	Florida, East	1000 ybp	1	Unknown
Grass Mesa	Colorado, West	1100-1400 ybp	1	(Breternitz, 1983)
Janey B. Goode	Illinois, Midwest	1000-1400 ybp	19	(Galloy, 2010)
Koster*	Illinois, Midwest	9000 ybp	3	(Morey and Wiant, 1992)
Little Bear	Alabama, East	3000-7000 ybp	1	(Webb and DeJarnette, 1948b)
Mayapan	Mexico	1000 ybp	4	(Masson and Lope, 2008)
McPhee Pueblo	Colorado, West	1100-1300 ybp	1	(Kane and Robinson, 1988)
Modoc	Missouri	9000 ybp	2	(Parmalee, 1959)
Perry	Alabama, East	3000-7000 ybp	3	(Webb and DeJarnette, 1942)
Prince Rupert Harbour	British Columbia, West	1500 ybp	5	(Stewart and Stewart, 1996)
Scioto Caverns	Missouri,	2000 ybp	7	(Potter and Baby,

Table 3.1 (cont'd)

	Midwest			1964)
Simonsen Bison Kill	Iowa, Midwest	7200-7600 ybp	1	(Agogino and Frankforter, 1960)
Yellow Jacket Pueblo	Colorado, West	800-1000 ybp	1	(Kuckelman, 2003)
Zhokhov	Eastern Siberia	8000 ybp	9	(Lee et al., 2015)

Table 3.2: A summary of sequencing results for each sample that was captured. The archaeological site the sample derives from is listed, as well as the percentage of sequencing reads that mapped to the mitogenome, the coverage of the mitogenome (with the maximum being 16727 bp), as well as the read depth, or the average number of reads that mapped to each individual base pair in the mitogenome sequence. The last column indicates whether or not this sequence was used in the analyses in this paper

Individual	Archaeological Site	% of reads that aligned	Mitogenome Coverage	Average Read Depth	Used in Analysis?
5MT23-16	Grass Mesa	5.53	16702	42.0256	Yes
5MT5-01	Yellow Jack Pueblo	37.01	16727	94.116	Yes
5MT4475-20	McPhee Pueblo	4.22	16190	7.34515	Yes
AM310A	Angel Mounds	30.68	16722	109.581	Yes
AM310B	Angel Mounds	9.98	16615	65.0508	Yes
AM310C	Angel Mounds	4.94	16531	34.3923	Yes
AM474	Angel Mounds	4.15	16709	50.3151	Yes
CAO1	Channel Islands	32.34	16727	131.411	Yes
CAW2	Channel Islands	28.45	16727	76.3536	Yes
CGG1	Zhokhov	21.64	16727	111.735	Yes
CGG2	Zhokhov	13.35	16727	79.41	Yes
CGG3	Zhokhov	26.38	16727	46.584	Yes
CGG4	Zhokhov	47.6	16727	100.533	Yes
CGG5	Zhokhov	45.91	16727	157.009	Yes
CGG6	Zhokhov	40.71	16727	63.6372	Yes
CGG7	Zhokhov	45.91	16727	168.146	Yes
CGG8	Zhokhov	11.45	16727	34.8966	Yes
CGG9	Zhokhov	36.65	16727	100.787	Yes
CGG10	Aachim Lighthouse	22.53	16727	31.3618	Yes
CGG11	Aachim Lighthouse	38.36	16727	115.82	Yes
CIAS	Channel Islands	9.79	16699	73.0302	Yes
CICVD	Channel Islands	1.36	16618	45.0662	Yes
CINH7	Channel Islands	14.62	16727	55.5992	Yes
CINHA	Channel Islands	59.05	16727	53.7035	Yes
CISG	Channel Islands	39.88	16727	116.642	Yes
CISNI4	Channel Islands	0.84	9862	1.805	No
Cox6	Cox	6	16532	26.0225	Yes
Fr11	Flint River	15.03	16422	15.1377	Yes
ISM21C	Anker	5.29	15713	5.32947	No
ISM070	Apple Creek	28.52	16561	57.6152	Yes
ISM090	Modoc	42.9	16715	100.653	Yes
ISM172	Simonsen Bison Kill	51.51	14903	7.60907	No
ISM256	Koster	80.92	16727	102.541	Yes
ISM357	Koster	13.95	16727	122.984	Yes

Table 3.2 (cont'd)

ISML50	Modoc	0.59	14527	5.52158	No
JBG1M	Janey B. Goode	0.1	16641	59.2071	Yes
JBG5	Janey B. Goode	26.85	16727	140.985	Yes
JBG11	Janey B. Goode	46.05	16727	181.116	Yes
JBG12	Janey B. Goode	29.47	16727	195.912	Yes
JBG13	Janey B. Goode	30.71	16727	195.719	Yes
JBG17	Janey B. Goode	50.64	16727	200.082	Yes
JBG19	Janey B. Goode	26.51	16727	151.944	Yes
JBG21	Janey B. Goode	11.93	16727	105.204	Yes
JBG24	Janey B. Goode	43.1	16727	194.479	Yes
JBG26	Janey B. Goode	43.05	16727	160.2	Yes
JBG32	Janey B. Goode	42.57	16727	179.076	Yes
JBG35	Janey B. Goode	34.71	16727	38.4495	Yes
JBG37	Janey B. Goode	39.79	16727	141.203	Yes
JBG41	Janey B. Goode	46.66	16727	187.745	Yes
JBG42	Janey B. Goode	46.78	16727	180.661	Yes
JBG43	Janey B. Goode	46.73	16727	201.066	Yes
JBG45	Janey B. Goode	37.88	16727	193.518	Yes
JBG48	Janey B. Goode	30.16	16727	169.955	Yes
JBG50	Janey B. Goode	29.6	16727	54.3204	Yes
LB2	Little Bear	8.63	16727	48.2802	Yes
May2	Mayapan	3.29	16719	66.8886	Yes
May3	Mayapan	3.51	16592	46.1615	Yes
May4	Mayapan	7.79	16623	71.7297	Yes
May10	Mayapan	9.42	16539	14.9061	Yes
OSU611	Scioto Caverns	13.98	16545	35.7524	Yes
OSU622	Scioto Caverns	16.54	16615	33.2042	Yes
OSU624	Scioto Caverns	30.68	16679	95.2518	Yes
OSU626	Scioto Caverns	60.24	16614	39.5706	Yes
OSU628	Scioto Caverns	74.11	16664	34.1894	Yes
OSU634	Scioto Caverns	0.41	16601	48.4206	Yes
OSU638	Scioto Caverns	39.96	16692	65.8408	Yes
P35	Perry	9.24	15793	12.5712	Yes
P39	Perry	1.6	9912	2.34897	No
P59	Perry	19.72	16727	31.3195	Yes
PRD1	Prince Rupert Harbour	28.73	16727	156.732	Yes
PRD9	Prince Rupert Harbour	33.08	16727	155.636	Yes
PRD10	Prince Rupert Harbour	0.03	16530	14.95	Yes
PRW5	Prince Rupert Harbour	5.9	13464	5.62366	No

Table 3.2 (cont'd)

PRW89	Prince Rupert Harbour	3.24	16065	8.71522	No

Table 3.3: A list of the modern dog haplotypes used for comparison to the ancient dogs and ancient and modern wolves. The haplotype is listed, as well as the Genbank accession number, the geographic origin (if known), and the reference the source can be found in.

Haplotype	Accession	Geographic Origin	Reference
A1	EU789697	SW Asia	(Pang et al., 2009)
A1b	EU789744	China	(Pang et al., 2009)
A2a	EU789675	China	(Pang et al., 2009)
A3a	EU789682	Thailand	(Pang et al., 2009)
A3b	EU789667	China	(Pang et al., 2009)
A3b	EU789693	China	(Pang et al., 2009)
A4	EU789669	China	(Pang et al., 2009)
A4a	EU789755	Japan	(Pang et al., 2009)
A4a	EU408262	USA	(Webb and Allard, 2009)
A5	EU789738	China	(Pang et al., 2009)
A5a	EU789663	China	(Pang et al., 2009)
A6	EU789688	China	(Pang et al., 2009)
A6a	EU789745	China	(Pang et al., 2009)
B1	EU789761	Korea	(Pang et al., 2009)
B2	EU789757	Kazakhstan	(Pang et al., 2009)
C1	EU789659	China	(Pang et al., 2009)
C1	EU789760	Europe	(Pang et al., 2009)
C1a	KM061488	Jamaica	(Duleba et al., 2015)
C2	KM061591	Fiji	(Duleba et al., 2015)
C2a	KJ637138	Belgium	(Verscheure et al., 2014)
D1	EU789654	Turkey	(Pang et al., 2009)
D1a	JF342859	USA	(Imes et al., 2012)
D1a	EU789655	Spain	(Pang et al., 2009)
D2	EU408288	USA	(Webb and Allard, 2009)
D2a	DQ480502	Scandinavia	(Björnerfeldt et al., 2006)
D2a	EU789656	Scandinavia	(Pang et al., 2009)
E	EU789662	Korea	(Pang et al., 2009)
F	AB499817	Japan	(Ishiguro and Nakano, 1996)

Table 3.4: A list of the modern dog haplogroup A haplotypes used to compare to the ancient dog mitogenomes. This list used as many mitogenomes from the USA as possible (ie if a haplotype was found in the US, that accession was used), and also tried to maximize geographic diversity by choosing regions with fewer dogs sampled where possible. The haplotype is listed, as well as the Genbank accession number, the geographic origin (if known), and the reference the source can be found in.

Haplotype	Accession	Geographic Origin	Reference
A1	AY656737	S. Africa	(Shahid et al., 2004)
A1a1	JF342810	USA	(Imes et al., 2012)
A1a1a	EU408259	USA	(Webb and Allard, 2009)
A1a1a1	KM061551	Russia	(Duleba et al., 2015)
A1a1a1a	JF342872	USA	(Imes et al., 2012)
A1a1a1a1	EU789720	Iran	(Pang et al., 2009)
A1a1a2	EU789733	Iran	(Pang et al., 2009)
A1a1a2a	EU408282	USA	(Webb and Allard, 2009)
A1a1a2b	KJ637044	Belgium	(Verscheure et al., 2014)
A1a1a3	EU408246	USA	(Webb and Allard, 2009)
A1a1a4	KM061527	France	(Duleba et al., 2015)
A1a1a5	EU408305	USA	(Webb and Allard, 2009)
A1a1a5a	AY656742	Great Britain	(Shahid et al., 2004)
A1a1a6	KM061478	Caucasus	(Duleba et al., 2015)
A1a1a7	KM061531	Poland	(Duleba et al., 2015)
A1a1a8	EU789780	Great Britain	(Björnerfeldt et al., 2006)
A1a1a9	KJ637036	Belgium	(Verscheure et al., 2014)
A1a1a10	KJ637042	Belgium	(Verscheure et al., 2014)
A1a1a10a	KJ637041	Belgium	(Verscheure et al., 2014)
A1a1a11	EU789709	France	(Pang et al., 2009)
A1a1a12	JF342822	USA	(Imes et al., 2012)
A1a1b	KM061588	Fiji	(Duleba et al., 2015)
A1a1b1	EU408286	USA	(Webb and Allard, 2009)
A1a1b1a	EU408249	USA	(Webb and Allard, 2009)
A1a1b1a1	KM061577	Tajikistan	(Duleba et al., 2015)
A1a1b1a2	KJ637054	Belgium	(Verscheure et al., 2014)
A1a1b1a3	KJ637056	Belgium	(Verscheure et al., 2014)
A1a1b1b	EU408295	USA	(Webb and Allard, 2009)
A1a1b2	JF342882	USA	(Imes et al., 2012)
A1a1c	EU789722	China	(Pang et al., 2009)
A1a1c1	EU789715	China	(Pang et al., 2009)
A1a1d	EU408264	USA	(Webb and Allard, 2009)
A1a1e	KM061479	Caucasus	(Duleba et al., 2015)
A1a1e1	KM061499	Costa Rica	(Duleba et al., 2015)
A1a2a	EU789768	Siberia	(Pang et al., 2009)

Table 3.4 (cont'd)

A1a2b	KJ637069	Belgium	(Verscheure et al., 2014)
A1b	EU789744	China	(Pang et al., 2009)
A1b1a1	EU408304	USA	(Webb and Allard, 2009)
A1b1a1a	EU408250	USA	(Webb and Allard, 2009)
A1b1a1a1	JF342904	USA	(Imes et al., 2012)
A1b1a1a2	KJ637073	Belgium	(Verscheure et al., 2014)
A1b1a1a3	EU408263	USA	(Webb and Allard, 2009)
A1b1a1b	KM061489	Jamaica	(Duleba et al., 2015)
A1b1a1c	KJ637079	Belgium	(Verscheure et al., 2014)
A1b1a2	EU408274	USA	(Webb and Allard, 2009)
A1b1a2a	AY656751	Great Britain	(Shahid et al., 2004)
A1b1a2a1	EU408245	USA	(Webb and Allard, 2009)
A1b1a2a2	KJ637097	Belgium	(Verscheure et al., 2014)
A1b1a2b	KM061528	France	(Duleba et al., 2015)
A1b1a2b1	JF342902	USA	(Imes et al., 2012)
A1b1a2c	KJ637090	Belgium	(Verscheure et al., 2014)
A1b1b	EU789759	DRC	(Pang et al., 2009)
A1b1b1a	JF342899	USA	(Imes et al., 2012)
A1b1b1b	KM061590	Fiji	(Duleba et al., 2015)
A1b2a1	EU408275	USA	(Webb and Allard, 2009)
A1b2a1a	JF342867	USA	(Webb and Allard, 2009)
A1b2a1a1	EU408276	USA	(Webb and Allard, 2009)
A1b2a1a2	KM061502	Russia	(Duleba et al., 2015)
A1b2a2	KM061529	France	(Duleba et al., 2015)
A1b2a2a	KM061523	France	(Duleba et al., 2015)
A1b2a2a1	JF342903	USA	(Imes et al., 2012)
A1b2a2b	EU789683	Great Britain	(Pang et al., 2009)
A1b2a2b1	KM061490	Costa Rica	(Duleba et al., 2015)
A1b2b	EU789689	India	(Pang et al., 2009)
A1b3	JF342885	USA	(Imes et al., 2012)
A1b3a	KM061589	Fiji	(Duleba et al., 2015)
A1b3a1	DQ480491	Great Britain	(Björnerfeldt et al., 2006)
A1b3b1	AY656741	Great Britain	(Shahid et al., 2004)
A1b3b2	EU789707	China	(Pang et al., 2009)
A1b3c	EU408289	USa	(Webb and Allard, 2009)
A1b4	EU789714	Spain	(Pang et al., 2009)
A1b5	EU789686	N Africa	(Pang et al., 2009)
A1b5a	EU789679	Saudi Arabia	(Pang et al., 2009)
A1b6	EU789717	China	(Pang et al., 2009)
A1b6a	EU789701	China	(Pang et al., 2009)

Table 3.4 (cont'd)

A1b7	EU789754	India	(Pang et al., 2009)
A1b7a	EU787767	India	(Pang et al., 2009)
A2a	EU789675	China	(Pang et al., 2009)
A2a1	EU789699	China	(Pang et al., 2009)
A2a1a	KM061542	Poland	(Duleba et al., 2015)
A2a1a1	EU789691	China	(Pang et al., 2009)
A2a1a1a	HM048871	China	(Li et al., 2008)
A2a1a1a1	EU789694	Tibet	(Pang et al., 2009)
A2a1b	EU789677	Thailand	(Pang et al., 2009)
A2a1b1	EU789685	China	(Pang et al., 2009)
A2b1	EU789674	China	(Pang et al., 2009)
A2b1a	EU789678	China	(Pang et al., 2009)
A2b1a1	EU789671	China	(Pang et al., 2009)
A2b2	EU789673	Thailand	(Pang et al., 2009)
A2b3	EU789681	China	(Pang et al., 2009)
A2b3a	DQ480499	Siberia	(Björnerfeldt et al., 2006)
A3a	EU789682	Thailand	(Pang et al., 2009)
A3a1	EU789676	China	(Pang et al., 2009)
A3b	EU789667	China	(Pang et al., 2009)
A3b1	EU789692	China	(Pang et al., 2009)
A4	EU789669	China	(Pang et al., 2009)
A4a	EU408262	USA	(Webb and Allard, 2009)
A5	EU789670	China	(Pang et al., 2009)
A5a	EU789670	China	(Pang et al., 2009)
A5a1	EU789665	China	(Pang et al., 2009)
A6	EU789688	China	(Pang et al., 2009)
A6a	EU408300	USA	(Webb and Allard, 2009)

Table 3.5: A list of all of the variants identified in the sample. The position refers to the corresponding position in the dog mitochondrial genome reference sequence, Genbank accession NC002008. The "Reference" column shows the base pair at that position in the reference sequence, while the "Variant" column shows the base pair variant found in one of the sequences mitogenomes. The individuals that were found to have that variant were listed next column. The type of mutation (Ts – transition, Tv – transversion, Del – deletion, Ins – insertion) is specified, along with the effect to the gene, if it is in a coding region. Amino acids are referred to using their standard three-letter abbreviation. The gene the nucleotide position is found in is also listed.

Position	Reference	Variant	Found in	Type	Effect on protein	Gene
165	T	C	CGG10, CGG11	Ts	None - RNA	12S rRNA
294	G	A	CAO1	Ts	None - RNA	12S rRNA
337	C	A	May3 May4 May10	Tv	None - RNA	12S rRNA
466	C	T	Fla	Ts	None - RNA	12S rRNA
526	T	C	CGG10, CGG11	Ts	None - RNA	12S rRNA
607	T	C	P35 LB2	Ts	None - RNA	12S rRNA
947	C	T	CGG10, CGG11	Ts	None - RNA	12S rRNA
966	C	T	CGG1	Ts	None - RNA	12S rRNA
1063	A	G	AM474	Ts	None - RNA	Val tRNA
1081	G	A	CINHA, PRD1, PRD9, PRD10	Ts	None - RNA	Val tRNA
1129	C	T	ISM070, OSU622, AM310A, 5MT501, Fla	Ts	None - RNA	16S rRNA
1403	G	A	AM310A	Ts	None - RNA	16S rRNA
1454	G	A	5MT501, CGG10, CGG11, CINHA, PRD1, PRD9, PRD10	Ts	None - RNA	16S rRNA
1485		A	5MT501, JBG42	In	None - RNA	16S rRNA
1522	G	A	Arg, May2, May3, May4, May10, P59	Ts	None - RNA	16S rRNA
1533	T	A	JBG13, JBG17	Tv	None - RNA	16S rRNA
1579	C	T	JBG19, JBG50	Ts	None - RNA	16S rRNA
1697	C	T	JBG5, JBG45	Ts	None - RNA	16S rRNA

Table 3.5 (cont'd)

1863	T	C	PRD1, PRD9, PRD10	Ts	None - RNA	16S rRNA
1977	G	A	CINHA, JBG42	Ts	None - RNA	16S rRNA
2176	A	G	Arg	Ts	None - RNA	16S rRNA
2184	T	C	JBG5, JBG45	Ts	None - RNA	16S rRNA
2198	T	C	OSU628	Ts	None - RNA	16S rRNA
2232	A	G	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	None - RNA	16S rRNA
2234	T	C	CGG7	Ts	None - RNA	16S rRNA
2401	G	A	AM310B, AM310C, Cox6, ISM070, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU624, OSU626, OSU628, OSU634, OSU638, P35	Ts	None - RNA	16S rRNA
2575	G	A	ISM070	Ts	None - RNA	16S rRNA
2578	A	G	ISM070	Ts	None - RNA	16S rRNA
2593	A	G	5MT501	Ts	None - RNA	16S rRNA
2631	A	G	Arg	Ts	None - RNA	16S rRNA

Table 3.5 (cont'd)

2678		G	AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10	Ins	None - RNA	Leu tRNA
2678		A	ISM070	Ins	None - RNA	Leu tRNA
2683	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	None - RNA	Leu tRNA
2779	C	T	CINHA	Ts	Silent	ND1
2800	C	T	CGG10, CGG11	Ts	Silent	ND1
2835	A	G	CGG7	Ts	Ile > Thr	ND1
2937	T	A	AM310B, AM310C, Cox6, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638	Tv	Glu > Val	ND1

Table 3.5 (cont'd)

2962	C	T	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ND1
2971	C	G	JBG37, JBG41, JBG42, JBG48	Tv	silent	ND1
3049	A	G	CGG1	Ts	Ser > Pro	ND1
3080	G	A	Fla, Kos, 5MT501, AM310A, ISM070, ISM256, ISM357	Ts	Arg > Trp	ND1
3127	C	T	5MT501	Ts	silent	ND1
3145	T	C	5MT520, CAO1, CAW2, CIAS, CISG	Ts	silent	ND1
3196	T	C	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, COx6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ND1
3203	G	A	10-May	Ts	silent	ND1
3250	C	T	2-May	Ts	Trp > Cys	ND1
3283	A	G	May3, May4, May10	Ts	silent	ND1
3287	C	T	CGG7	Ts	Asp > Asn	ND1
3388	G	A	2-May	Ts	silent	ND1

Table 3.5 (cont'd)

3406	C	T	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, COx6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ND1
3514	T	C	LB2	Ts	silent	ND1
3598	G	A	JBG19, JBG50	Ts	silent	ND1
3662	G	A	Fla	Ts	Arg > trp	ND1
3670	C	T	2-May	Ts	silent	ND1
3724	G	A	AM474, JBG13, JBG17	Ts	None - RNA	Ile tRNA
3903	T	C	JBG13, JBG17	Ts	None - RNA	Met tRNA
3944	G	A	JBG12, JBG35	Ts	Arg > trp	ND2
4070	A	T	Arg	Tv	Tyr > Asn	ND2
4150	T	C	JBG13, JBG17	Ts	silent	ND2
4200	T	C	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59	Ts	Met > Thr	ND2
4212	C	T	JBG13, JBG17	Ts	Thr > Ile	ND2
4291	A	G	CAW2	Ts	silent	ND2
4333	C	T	AM310B, AM310C, JBG12, JBG32, JBG35	Ts	silent	ND2
4375	T	C	JBG24, JBG41	Ts	silent	ND2
4390	T	A	Arg	Tv	Glu > Asp	ND2
4391	A	G	AM474	Ts	Cys >	ND2

Table 3.5 (cont'd)

					Gly	
4468	C	T	CGG2	Ts	silent	ND2
4511	A	G	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638	Ts	Tyr > His	ND2
4517	G	A	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638	Ts	Val > Ile	ND2
4588	C	T	P59	Ts	Trp > Cys	ND2
4591	G	A	AM310B, AM310C, CGG7, Cox6, Fr11, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35	Ts	silent	ND2
4726	G	A	CGG7	Ts	silent	ND2
4732	T	C	JBG11, JBG21, JBG26, JBG43	Ts	silent	ND2
4759	G	A	CINHA, PRD1, PRD9, PRD10	Ts	silent	ND2
4780	C	T	CGG10, CGG11	Ts	silent	ND2
4810	C	T	PRD1, PRD9, PRD10	Ts	silent	ND2
4940	T	C	Fla	Ts	Asn > Asp	ND2
4949	T	C	CINHA, PRD1, PRD9, PRD10	Ts	Asn > Asp	ND2
5126	C	T	CGG10, CGG11	Ts	None - RNA	Asn tRNA
5162	T	C	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59	Ts	None - RNA	Asn tRNA

Table 3.5 (cont'd)

5259	A	G	Arg, 5MT316, 5MT520, AM474, CAO1, CAW2, CIAS, CISG, JBG13, JBG17, May2, May3, May4, May10	Ts	None - RNA	Cys tRNA
5272	A	G	CAO1, CAW2, CIAS	Ts	None - RNA	Cys tRNA
5279	T	C	P59	Ts	None - RNA	Cys tRNA
5366	A	G	JBG5	Ts	silent	COI
5367	C	T	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	Asp > asn	COI
5444	T	C	5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG2, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10		silent	COI
5498	C	T	AM474	Ts	silent	COI
5564	C	T	CGG3, CGG5	Ts	silent	COI
5567	C	T	CGG6	Ts	silent (stop to stop?)	COI
5597	G	A	CGG3, CGG5	Ts	silent	COI
5699	G	A	Arg, May2, May3, May4, May10	Ts	silent	COI

Table 3.5 (cont'd)

5744	G	A	CGG3, CGG5	Ts	silent	COI
6053	C	T	CINHA	Ts	silent	COI
6065	A	G	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	COI
6257	G	A	5MT501	Ts	silent	COI
6407	A	G	CGG10, CGG11	Ts	silent	COI
6506	T	C	May2, May3, May4, May10	Ts	silent	COI
6530	C	T	OSU628	Ts	silent	COI
6533	C	T	CGG5	Ts	silent	COI
6554	T	C	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	COI
6557	T	C	CICVD	Ts	silent	COI

Table 3.5 (cont'd)

6572	T	C	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	COI
6662	T	A	5MT501	Tv	silent	COI
6674	C	T	CGG3, CGG5	Ts	silent	COI
6749	T	A	CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9	Tv	Glu > Asp	COI
6798	A	G	AM474	Ts	silent	COI
6854	C	T	OSU638	Ts	silent	COI
6882	A	G	AM310A, OSU628	Ts	Stop > Gln?	COI
6959	C	T	P59	Ts	None - RNA	Ser tRNA
7014		G	CAO1, CAW2, CIAS	Ins	None - RNA	Asp tRNA
7081	T	C	Fla, Arg, Kos, 5MT501, 5MT316, 5MT520, AM310A, AM474, CAO1, CAW2, CIAS, CISG, ISM070, ISM256, ISM357, JBG1, JBG13, JBG17, May2, May3, May4, May10	Ts	silent	COII
7145	G	A	JBG11, JBG21, JBG26, JBG37, JBG42, JBG43, JBG48	Ts	Val > Ile	COII
7166	A	G	Fla	Ts	Tyr > His	COII
7186	C	A	Arg	Tv	Trp > Cys	COII
7264	T	C	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59	Ts	silent	COII

Table 3.5 (cont'd)

7291	A	G	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	COII
7492	T	C	P35 LB2	Ts	silent	COII
7493	A	G	CICVD	Ts	Ile > Val	COII
7526	G	A	CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9	Ts	Val > Ile	COII
7552	C	T	Arg, May2, May3, May4, May10	Ts	silent	COII
7585	T	C	Arg, 5MT316, 5MT520, AM474, CAO1, CAW2, CIAS, CISG, JBG13, JBG17, May2, May3, May4, May10	Ts	silent	COII
7618	G	A	5MT520	Ts	silent	COII
7727	G	A	CGG10, CGG11	Ts	Non-coding	Non-coding
7729	T	C	ISM070	Ts	Non-coding	Non-coding
7838	A	G	CINHA	Ts	silent	ATPase subunit 8
7851	T	C	CGG2	Ts	Lys > Glu	ATPase subunit 8
7914	A	G	CGG10, CGG11	Ts	Tyr > His	ATPase subunit 8
7918	C	T	P59	Ts	Thr > Ile	ATPase subunit 8
8048	A	G	CGG10, CGG11	Ts	Ile > Val	ATPase subunit 6
8107	A	G	FR11	Ts	silent	ATPase subunit 6

Table 3.5 (cont'd)

8224	A	G	JBG11, JBG21, JBG26, JBG37, JBG42, JBG43, JBG48	Ts	silent	ATPase subunit 6
8281	T	C	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ATPase subunit 6
8317	C	T	AM474	Ts	silent	ATPase subunit 6
8323	A	G	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ATPase subunit 6
8331	C	T	May3, May4, May10	Ts	Arg > His	ATPase subunit 6
8335	C	T	5MT520, AM474, CAO1, CAW2, CIAS, CISG, JBG13, JBG17	Ts	silent	ATPase subunit 6
8368	C	T	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2,	Ts	silent	ATPase subunit 6

Table 3.5 (cont'd)

			CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10			
8411	T	C	CGG10, CGG11	Ts	Lys > Glu	ATPase subunit 6
8419	A	G	CINHA, PRD1, PRD9, PRD10	Ts	silent	ATPase subunit 6
8525	A	G	Fla	Ts	Cys > Arg	ATPase subunit 6
8553	T	C	CGG3, CGG5, Fr11	Ts	Ile > Thr	ATPase subunit 6
8558	C	T	PRD9	Ts	Glu > Lys	ATPase subunit 6
8703	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	COIII
8742	A	G	LB2	Ts	silent	COIII
8754	T	C	Arg	Ts	silent	COIII
8757	C	T	JBG32	Ts	silent	COIII

Table 3.5 (cont'd)

8764	G	T	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Tv	silent	COIII
8765	C	T	Kos, ISM256	Ts	Arg > Gln	COIII
8782	T	C	CGG7, CGG8	Ts	Lys > Glu	COIII
8802	C	T	JBG24, JBG41	Ts		COIII
8807	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	Cys	Tyr
8894	T	C	JBG32	Ts	Gln > Arg	COIII
8904	C	T	PRD9	Ts	silent	COIII
9072	T	C	PRD9	Ts	silent	COIII
9165	G	A	CGG2, CGG6, CGG9	Ts	silent	COIII
9193	A	G	ISM070	Ts	Cys > Arg	COIII
9270	T	C	2-May	Ts	silent (stop > stop)	COIII

Table 3.5 (cont'd)

9324	C	T	Fla, Kos, 5MT501, AM310A, ISM070, ISM256, ISM357	Ts	silent	COIII
9366	T	C	ISM070	Ts	silent	COIII
9585	T	C	PRD9	Ts	silent	ND3
9630	A	G	Arg	Ts	silent	ND3
9831	A	G	AM310A, OSU622	Ts	silent	ND3
9860		A	5MT316, 5MT501, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ins	None - RNA	Arg tRNA
9919	G	A	Cox6	Ts	silent	ND4L
9928	T	C	JBG13, JBG17	Ts	silent	ND4L
10056	T	C	Arg	Ts	silent	ND4L
10111	T	C	P59	Ts	silent	ND4L
10263	C	T	CICVD, CINH7, CINHA, PRD1, PRD9, PRD10	Ts		ND4
10281	G	A	Arg, 5MT316, 5MT520, AM474, CAO1, CAW2, CIAS, CISG, JBG13, JBG17, May2, May3, May4, May10	Ts	silent	ND4
10285	A	G	AM310A	Ts	Cys > Arg	ND4
10318	T	C	AM310A	Ts	Lys > Glu	ND4
10344	T	C	OSU624, OSU626, OSU634	Ts	silent	ND4
10354	T	C	CGG10, CGG11	Ts	Arg > Gly	ND4
10467	A	G	AM310B, AM310C	Ts	silent	ND4
10485	T	C	AM310B, AM310C, JBG12, JBG32, JBG35, OSU638	Ts	silent	ND4

Table 3.5 (cont'd)

10512	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ND4
10689	T	C	P35	Ts	silent	ND4
10698	T	C	ISM090	Ts	silent	ND4
10735	A	G	AM310B, AM310C, JBG12, JBG32, JBG35	Ts	Ile > Val	ND4
10863	A	G	AM310A	Ts	silent	ND4
10917	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10	Ts	silent	ND4
10936	A	G	5MT501	Ts	Ile > Val	ND4

Table 3.5 (cont'd)

10992	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	silent	ND4
11082	G	A	JBG13, JBG17	Ts	silent	ND4
11115	A	G	CGG7	Ts	silent	ND4
11227	A	G	May2, May3, May4, May10	Ts	Ile > Val*	ND4
11229	T	C	May2, May3, May4, May10	Ts	Ile > Val*	ND4
11265	A	G	CGG3, CGG5	Ts	silent	ND4
11310	T	C	CGG1	Ts	silent	ND4
11364	A	G	JBG5, JBG45	Ts	silent	ND4
11372	T	C	Arg, May2, May3, May4, May10	Ts	silent	ND4
11393	T	C	AM310A	Ts	silent	ND4
11577	C	T	OSU611	Ts	silent	ND4
11613	G	A	JBG13, JBG17	Ts	None - RNA	His tRNA
11727	A	G	CICVD, CINH7	Ts	None - RNA	Leu tRNA
11732	G	A	CGG2, CGG4, CGG6, CGG7, CGG8, CGG9	Ts	None - RNA	Leu tRNA
11816	T	C	PRD9	Ts	silent	ND5
11828	G	A	Arg	Ts	silent	ND5
11846	T	C	Arg, May2, May3, May4, May10	Ts	silent	ND5

Table 3.5 (cont'd)

11892	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	Ala > Thr	ND5
11950	G	A	CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9	Ts	Pro > Leu	ND5
11954	G	A	5MT501	Ts	silent	ND5
12090	A	G	FR11	Ts	Tyr > His	ND5
12164	A	G	P35	Ts	silent	ND5
12388	C	T	CICVD, CINH7, CINHA, PRD1, PRD9, PRD10	Ts	Cys > Tyr	ND5
12398	T	C	ISM070	Ts	silent	ND5
12429	C	T	Arg	Ts	Asp > Asn	ND5
12593	T	C	OSU638	Ts	Ser > Arg	ND5
12620	G	A	PRD1, PRD9, PRD10	Ts	silent	ND5
12728	T	C	Cox6	Ts	silent	ND5
12788	T	C	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10	Ts	silent	ND5

Table 3.5 (cont'd)

12831	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10	Ts	silent	ND5
12959	G	A	CGG3, CGG5	Ts	silent	ND5
13035	G	A	Arg	Ts	Ala > Thr	ND5
13046	C	T	CINHA, PRD1, PRD9, PRD10	Ts	silent	ND5
13118	T	C	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59	Ts	silent	ND5
13168	C	T	5MT501	Ts	Arg > His	ND5
13181	C	T	AM310B, AM310C, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2	Ts	silent	ND5
13202	T	C	CGG10, CGG11	Ts	silent	ND5
13297	C	A	Arg, 5MT520, AM474, CAO1, CAW2, CIAS, CISG, JBG13, JBG17, May2, May3, May4, May10	Tv	Trp > Leu	ND5

Table 3.5 (cont'd)

13299	T	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10	Tv	Ser > Cys	ND5
13319	C	T	Fla	Ts	Leu > Phe	ND5
13322	T	C	CGG3, CGG5	Ts	silent	ND5
13415	G	A	AM310B, AM310C, Cox6, FR11, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638	Ts	silent	ND5
13427	T	C	CGG10, CGG11	Ts	silent	ND5
13444	T	C	CAO1, CAW2, CIAS	Ts	Ile > Thr	ND5
13459	T	C	CGG10, CGG11	Ts	Ile > Thr	ND5
13481	T	C	JBG13, JBG17	Ts	silent	ND5
13572	C	T	CGG1, CGG2, CGG4, CGG6, CGG7, CGG8, CGG9	Ts	Val > Leu	ND5
13771	C	T	AM310A	Ts	silent	ND6
13799	G	A	Kos, ISM256	Ts	Ala > Val	ND6
13828	A	G	May3, May4, May10	Ts	Val > Ile	ND6
13864	C	T	CGG10, CGG11	Ts	silent	ND6
13884	T	C	AM310B, AM310C, Cox6, ISM090, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU626, OSU628, OSU634, OSU638, P35	Ts	silent	ND6
13980	C	T	Cox6	Ts	Val > Ile	ND6
13995	T	C	CGG3, CGG5	Ts	silent	ND6

Table 3.5 (cont'd)

14019	G	A	5MT501, 5MT520, AM310A, ISM070	Ts	silent	ND6
14023	A	G	P59	Ts	silent	ND6
14149	A	G	PRD9	Ts	None - RNA	Glu tRNA
14230	C	T	CAO1, CAW2, CIAS	Ts	silent	cytB
14288	T	C	PRD1, PRD10	Ts	Asn > Asp	cytB
14315	A	G	CAW2	Ts	Stop > Gln?	cytB
14329	A	G	CGG10, CGG11	Ts	silent	cytB
14368	T	C	JBG13, JBG17	Ts	silent	cytB
14383	C	G	CGG10, CGG11	Tv	Trp > Cys	cytB
14401	T	C	OSU624, OSU626	Ts	silent	cytB
14425	T	C	CGG10, CGG11	Ts	silent	cytB
14467	C	T	CGG10, CGG11, JBG11, JBG21, JBG26, JBG37, JBG42, JBG43, JBG48	Ts	silent	cytB
14475	T	C	Arg	Ts	His > Arg	cytB
14525	A	G	CICVD, CINH7	Ts	Stop > Gln?	cytB
14543	T	C	ISM090	Ts	Lys > Glu	cytB
14716	C	T	Fla	Ts	silent	cytB
14728	T	C	CAO1, CAW2, CIAS, May2, May3, May4, May10	Ts	silent	cytB
14764	A	G	Fla	Ts	silent	cytB
14788	A	G	Arg, May2, May3, May4, May10	Ts	silent	cytB
14900	A	G	JBG5	Ts	Tyr > His	cytB
14911	T	C	May3, May4	Ts	silent	cytB
14970	A	T	CGG3, CGG5	Tv	Asn > Ile	cytB
14998	A	G	CGG10, CGG11	Ts	silent	cytB
15052	T	C	AM474	Ts	silent	cytB
15214	G	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37,	Ts	silent	cytB

Table 3.5 (cont'd)

			JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10			
15232	C	T	JBG5	Ts	silent	cytB
15280	A	G	May3, May4, May10	Ts	silent	cytB
15284	A	G	CINHA, PRD1, PRD9, PRD10	Ts	Cys > Arg	cytB
15294	T	C	CINHA	Ts	Gln > Arg	cytB
15338	T	C	5MT520, CAO1, CAW2, CIAS	Ts	None - RNA	Thr tRNA
15372	G	A	CICVD, CINH7	Ts	None - RNA	Thr tRNA
15378	G	A	AM474, JBG13, JBG17	Ts	None - RNA	Thr tRNA
15484	A	G	AM474, CAO1, CAW2, CIAS, CISG, JBG13, JBG17	Ts	Non-coding	control region
15553	A	G	CINHA, PRD1, PRD9, PRD10	Ts	Non-coding	control region
15621	C	T	May2, May3, May4, May10	Ts	Non-coding	control region
15625	T	C	5MT501, LB2, May3, May4	Ts	Non-coding	control region
15627	A	G	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10	Ts	Non-coding	control region
15639	T	A	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA,	Tv	Non-coding	control region

Table 3.5 (cont'd)

			CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P59, PRD1, PRD9, PRD10			
15639	T	G	P59	Tv	Non-coding	control region
15643	A	G	CGG2	Ts	Non-coding	control region
15651	C	T	May2, May3, May4, May10	Ts	Non-coding	control region
15652	G	A	CGG10, CGG11, May2	Ts	Non-coding	control region
15665	T	C	Arg	Ts	Non-coding	control region
15750	C	T	AM310A	Ts	Non-coding	control region
15807	C	T	5MT520, AM474, CISG, JBG13, JBG17, May2, May3, May4, May10	Ts	Non-coding	control region
15813	C	T	AM474, JBG13, JBG17	Ts	Non-coding	control region
15814	C	T	Fla, Kos, Arg, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	Non-coding	control region
15912	C	T	5MT501	Ts	Non-coding	control region
15931	A		CGG2, CICVD, CINH7, CINHA, PRD9, PRD10	Del	Non-coding	control region
15931	A	G	PRD1	Ts	Non-	control

Table 3.5 (cont'd)

					coding	region
15931	G		CGG2, PRD10	Del	Non-coding	control region
15955	C	T	CINHA, May3, May4	Ts	Non-coding	control region
16003	A	G	AM310A	Ts	Non-coding	control region
16025	T	C	5MT501, AM310A, AM474, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CINHA, CISG, ISM070, ISM256, ISM357, JBG13, JBG17, May2, May3, May4, May10, PRD1, PRD9, PRD10	Ts	Non-coding	control region
16032	A	G	CGG6		Non-coding	control region
16148	A	G	AM310A, AM310B, AM310C, CGG4, CIAS, CICVD, ISM256, JBG12, JBG24, JBG35,	Ts	Non-coding	control region
16158	A	G	5MT520, JBG32, JBG42		Non-coding	control region
16168	A	G	5MT520, CGG4, CGG7, JBG24, JBG48		Non-coding	control region
16178	A	G	CGG9		Non-coding	control region
16188	G	A	CGG9, JBG24		Non-coding	control region
16216	G	A	CGG5		Non-coding	control region
16226	G	A	CGG5		Non-coding	control region
16236	G	A	CGG5		Non-coding	control region
16256	G	A	CGG5		Non-coding	control region
16276	G	A	CGG5		Non-coding	control region
16288	A	G	CGG7		Non-coding	control region
16338	G	A	JBG12		Non-coding	control region
16378	G	A	CGG5		Non-coding	control region
16388	A	G	CGG1, CGG2, CGG5, JBG12, JBG13, JBG41, JBG42, JBG45, JBG48	Ts	Non-coding	control region

Table 3.5 (cont'd)

16398	A	G	CGG1, CGG2, CGG3, CIAS, ISM090, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, OSU628, P35	Ts	Non-coding	control region
16408	G	A	AM310A, AM310B, AM310C, CGG4, CGG5, CGG6, CGG7, JBG17, OSU624, OSU626	Ts	Non-coding	control region
16418	A	G	5MT316, 5MT501, AM310B, AM310C, AM474, CAO1, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CINH7, CINHA, CISG, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, PRD1, PRD9	Ts	Non-coding	control region
16431	C	T	5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAO1, CAW2, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CGG10, CGG11, CIAS, CICVD, CINH7, CINHA, CISG, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG12, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU624, OSU626, OSU628, OSU634, OSU638, P35, P59, PRD1, PRD9, PRD10	Ts	Non-coding	control region
16436	G	A	May2, May3, May10, OSU611	Ts	Non-coding	control region
16539	C	T	ISM070	Ts	Non-coding	control region
16610	C	T	Am310B, AM310C, Cox6, JBG5, JBG11, JBG12, JBG19, JBG21, JBG24, JBG26, JBG32, JBG35, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, OSU611, OSU624, OSU626, OSU628, OSU634, OSU638, P35	Ts	Non-coding	control region

Table 3.5 (cont'd)

16671	T	C	Fla, Kos, 5MT316, 5MT501, 5MT520, AM310A, AM310B, AM310C, AM474, CAW2, CGG1, CGG2, CGG3, CGG4, CGG5, CGG6, CGG7, CGG8, CGG9, CIAS, CICVD, CINH7, CINHA, CISG, Cox6, FR11, ISM070, ISM090, ISM256, ISM357, JBG1, JBG5, JBG11, JBG13, JBG17, JBG19, JBG21, JBG24, JBG26, JBG32, JBG37, JBG41, JBG42, JBG43, JBG45, JBG48, JBG50, LB2, May2, May3, May4, May10, OSU611, OSU622, OSU626, OSU634, OSU628, OSU638, P59, PRD1, PRD9, PRD10	Ts	Non-coding	control region
16671		C	JBG12	Ins	Non-coding	control region
16672	C	T	CGG10, CGG11	Ts	Non-coding	control region

Table 3.6: Genetic diversity measures for all dog populations with N>1

Archaeological Site	Sample Size	Nucleotide Diversity	N. Diversity s.d.	Theta S	Theta S s.d.
Angel Mounds	4	.001676	.001120	26.182	14.413
Channel Islands	7	.001066	.000620	17.143	8.701
Janey B. Goode	19	.000800	.000422	12.889	6.077
Koster	3	.000126	.000119	1.333	1.098
Mayapan	4	.000425	.000323	7.636	4.447
Perry	2	.000924	.000956	14.000	10.247
Prince Rupert Harbour	3	.000372	.000304	5.333	3.528
Scioto	7	.000239	.000157	5.306	2.709
Siberia	11	.001217	.000659	20.485	8.374
Southwest	3	.001034	.000799	13.333	8.319

Table 3.7: The partitions of the mitogenome alignment as identified by PartitionFinder 2.0, for use in BEAST. The alignment was partitioned by gene and each gene was analyzed separately. These numbers correspond to the position in the alignment used for the BEAST analysis. In the suggested model section, G indicates gamma, I indicates invariant sites, and X indicates that the base frequencies should be estimated, rather than using the empirical values.

Partition	Range of base pairs	Suggested Model
1	1-1024, 1093-2753, 3710-3775, 4964-5114, 5187-5223, 6903-7043, 7745-7812, 9438-9505, 9852-9921, 11592-11790, 14123-14195	HKY+G+X
2	1025-1092. 3852-3921, 5224-5288, 5358-6902, 15471-16434	HKY+I+G
3	2754-3709, 3776-3851, 3922-4963, 5115-5186, 5289-5357, 7044-7744, 7813-9437, 9506-9851, 9922-11591, 11791-14122, 14196-15470	HKY+I+X

Table 3.8: A list of all published dog mitogenomes that are part of haplogroup A4, which is the modern dog haplogroup that is closest to ancient American dogs. Geographic origin and breed are listed as specifically as possible, and both are taken from the original manuscript the sequence appeared in.

Accession	Haplotype	Paper	Geographic Origin	Breed
EU408262	A4a	(Webb and Allard, 2009)	USA	Chihuahua
EU789669	A4	(Pang et al., 2009)	Shanxixian, China	Unknown
EU789755	A4a	(Pang et al., 2009)	Japan	Japanese Spitz
KM061562	A4a	(Duleba et al., 2015)	Russia	Pekingese
EU789664	A4	(Pang et al., 2009)	Laem Ngop,Thailand	Unknown
KF002306	A4	(Angleby et al., 2014)	Tibet	Shih tzu

References

Achilli, A., Perego, U.A., Bravi, C.M., Coble, M.D., Kong, Q.-P., Woodward, S.R., Salas, A., Torroni, A., Bandelt, H.-J., 2008. The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. PloS ONE. 3, e1764.

Achilli, A., Perego, U.A., Lancioni, H., Olivieri, A., Gandini, F., Hooshiar Kashani, B., Battaglia, V., Grugni, V., Angerhofer, N., Rogers, M.P., Herrera, R.J., Woodward, S.R., Labuda, D., Smith, D.G., Cybulski, J.S., Semino, O., Malhi, R.S., Torroni, A., 2013. Reconciling migration models to the Americas with the variation of North American native mitogenomes. Proceedings of the National Academy of Sciences of the United States of America. 110, 14308–13.

Acosta, A., Loponte, D., Garcia Esponda, C., 2011. Primer registro de perro doméstico prehispánico (Canis familiaris) entre los grupos cazadores recolectores del humedal de Paraná inferior (Argentina). Antípoda Revista de Antropología y Arqueología. 13, 175–199.

Agogino, G.A., Frankforter, W.D., 1960. A Paleo-Indian Bison-Kill in Northwestern Iowa. American Antiquity. 25, 414–415.

Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., Malaspinas, A., Margaryan, A., Higham, T., Chivall, D., Lynnerup, N., Harvig, L., Baron, J., Casa, P. Della, Dąbrowski, P., Duffy, P.R., Ebel, A. V., Epimakhov, A., Frei, K., Furmanek, M., Gralak, T., Gromov, A., Gronkiewicz, S., Grupe, G., Hajdu, T., Jarysz, R., Khartanovich, V., Khokhlov, A., Kiss, V., Kolář, J., Kriiska, A., Lasak, I., Longhi, C., McGlynn, G., Merkevicius, A., Merkyte, I., Metspalu, M., Mkrtchyan, R., Moiseyev, V., Paja, L., Pálfi, G., Pokutta, D., Pospieszny, Ł., Price, T.D., Saag, L., Sablin, M., Shishlina, N., Smrčka, V., Soenov, V.I., Szeverényi, V., Tóth, G., Trifanova, S. V., Varul, L., Vicze, M., Yepiskoposyan, L., Zhitenev, V., Orlando, L., Sicheritz-Pontén, T., Brunak, S., Nielsen, R., Kristiansen, K., Willerslev, E., 2015. Population genomics of Bronze Age Eurasia. Nature. 522, 167–172.

Allitt, S., Stewart, R.M., Messner, T., 2009. The Utility of Dog Bone (*Canis Familiaris*) in Stable Isotope Studies for Investigating the Presence of Prehistoric Maize (*Zea mays ssp. mays*): A Preliminary Study. North American Archaeologist. 29, 343–367.

Alt, S.M., 2010. Complexity in action(s): Retelling the Cahokia story. In: Ancient Complexities: New Perspectives on Pre-Columbian North America. pp. 119–137.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology. 215, 403–10.

Angleby, H., Oskarsson, M., Pang, J., Zhang, Y.P., Leitner, T., Braham, C., Arvestad, L., Lundeberg, J., Webb, K.M., Savolainen, P., 2014. Forensic informativity of ~3000 bp of coding sequence of domestic dog mtDNA. Journal of Forensic Sciences. 59, 898–908.

Bandelt, H.-J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution. 16, 37–48.

Battaglia, V., Grugni, V., Perego, U.A., Angerhofer, N., Gomez-Palmieri, J.E., Woodward, S.R., Achilli, A., Myres, N., Torroni, A., Semino, O., 2013. The first peopling of South America: new evidence from Y-

chromosome haplogroup Q. PloS ONE. 8, e71390.

Björnerfeldt, S., Webster, M.T., Vilà, C., 2006. Relaxation of selective constraint on dog mitochondrial DNA following domestication. Genome research. 16, 990–4.

Bluhm, E.A., Liss, A., 1961. Anker Site. In: Chicago Area Archaeology. Illinois Archaeological Survey, Urbana.

Bodner, M., Perego, U. A., Huber, G., Fendt, L., Röck, A.W., Zimmermann, B., Olivieri, A., Gómez-Carballa, A., Lancioni, H., Angerhofer, N., Bobillo, M.C., Corach, D., Woodward, S.R., Salas, A., Achilli, A., Torroni, A., Bandelt, H.-J., Parson, W., 2012. Rapid coastal spread of First Americans: novel insights from South America's Southern Cone mitochondrial genomes. Genome Research. 22, 811–20.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology. 10, 1–6.

Boyko, A.R., Boyko, R.H., Boyko, C.M., Parker, H.G., Castelhano, M., Corey, L., Degenhardt, J.D., Auton, A., Hedimbi, M., Kityo, R., Ostrander, E. a, Schoenebeck, J., Todhunter, R.J., Jones, P., Bustamante, C.D., 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. Proceedings of the National Academy of Sciences of the United States of America. 106, 13903–8.

Breternitz, D.A., 1983. Dolores Archaeological Program: Field Investigations and Analysis - 1978. Denver.

Brose, D.S., Greber, N., 1979. Hopewell archaeology. Kent State University Press.

Brown, J.A., Kerber, R.A., Winters, H.D., 1990. Trade and the Evolution of Exchange Relations at the Beginning of the Mississippian Period. In: Smith, B.D. (Ed.), The Mississippian Emergence. pp. 251–280.

Brown, S.K., Darwent, C.M., Sacks, B.N., 2013. Ancient DNA evidence for genetic continuity in arctic dogs. Journal of Archaeological Science. 40, 1279–1288.

Brown, S.K., Pedersen, N.C., Jafarishorijeh, S., Bannasch, D.L., Ahrens, K.D., Wu, J.-T., Okon, M., Sacks, B.N., 2011. Phylogenetic distinctiveness of Middle Eastern and Southeast Asian village dog Y chromosomes illuminates dog origins. PloS ONE. 6, e28496.

Brown, W.M., George, M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences of the United States of America. 76, 1967–1971.

Cantwell, A., 1980. Middle Woodland dog ceremonialism in Illinois. The Wisconsin Archeologist. 61, 480–496.

Castroviejo-Fisher, S., Skoglund, P., Valadez, R., Vilà, C., Leonard, J. A., 2011. Vanishing native American dog lineages. BMC Evolutionary Biology. 11, 73.

Charles, D.K., Buikstra, J., 2006. Recreating Hopewell. University Press of Florida.

Clutton-Brock, J., Noe-Nygaard, N., 1990. New Osteological and C-Isotope Evidence on Mesolithic Dogs : Companions to Hunters and Fishers at Star Carr , Seamer Carr and Kongemose. Journal of Archaeological Science. 17, 643–653.

Cooper, A., Poinar, H., 2000. Ancient DNA: do it right or not at all. Science. 289, 10–11.

Deguilloux, M.F., Moquel, J., Pemonge, M.H., Colombeau, G., 2009. Ancient DNA supports lineage replacement in European dog gene pool: insight into Neolithic southeast France. Journal of Archaeological Science. 36, 513–519.

Dillehay, T., Collins, M., 1988. Early cultural evidence from Monte Verde in Chile. Nature. 332, 150–152.

Ding, Z.-L., Oskarsson, M., Ardalan, a, Angleby, H., Dahlgren, L.-G., Tepeli, C., Kirkness, E., Savolainen, P., Zhang, Y.-P., 2012. Origins of domestic dog in southern East Asia is supported by analysis of Y-chromosome DNA. Heredity. 108, 507–14.

Drake, A.G., Coquerelle, M., Colombeau, G., 2015. 3D morphometric analysis of fossil canid skulls contradicts the suggested domestication of dogs during the late Paleolithic. Scientific Reports. 5, 1–8.

Drummond, A.J., Suchard, M. a, Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 29, 1969–73.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution. 22, 1185–92.

Druzhkova, A.S., Thalmann, O., Trifonov, V. A., Leonard, J. A., Vorobieva, N. V., Ovodov, N.D., Graphodatsky, A.S., Wayne, R.K., 2013. Ancient DNA Analysis Affirms the Canid from Altai as a Primitive Dog. PLoS ONE. 8, e57754.

Duchêne, S., Archer, F.I., Vilstrup, J., Caballero, S., Morin, P. A., 2011. Mitogenome phylogenetics: The impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PLoS ONE. 6, e27138.

Duleba, A., Skonieczna, K., Bogdanowicz, W., Malyarchuk, B., Grzybowski, T., 2015. Complete mitochondrial genome database and standardized classification system for Canis lupus familiaris. Forensic Science International: Genetics. 19, 123–129.

Dulik, M.C., Zhadanov, S.I., Osipova, L.P., Askapuli, A., Gau, L., Gokcumen, O., Rubinstein, S., Schurr, T.G., 2012. Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between Native Americans and Indigenous Altaians. American Journal of Human Genetics. 90, 229–46.

Emerson, T.E., 1997. Cahokia and the Archaeology of Power. The University of Alabama Press, Tuscaloosa, Alabama.

Emerson, T.E., Lewis, R.B., 1999. Cahokia and the hinterlands: middle Mississippian cultures of the Midwest. University of Illinois Press.

Emslie, S., 1978. Dog Burials from Mancos Canyon, Colorado. The Kiva. 43, 167–182.

Eshleman, J.A., Malhi, R.S., Johnson, J.R., Kaestle, F.A., Lorenz, J., Smith, D.G., 2004. Mitochondrial DNA and prehistoric settlements: native migrations on the western edge of North America. Human Biology. 76, 55–75.

Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10, 564–567.

Fix, A.G., 2005. Rapid deployment of the five founding amerind mtDNA haplogroups via coastal and riverine colonization. American Journal of Physical Anthropology. 128, 430–436.

Forster, P., Harding, R., Torroni, a, Bandelt, H.J., 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. American Journal of Human Genetics. 59, 935–45.

Frantz, L.A.F., Mullin, V.E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., Linderholm, A., Mattiangeli, V., Teasdale, M.D., Dimopoulos, E.A., Tresset, A., Duffraisse, M., McCormick, F., Bartosiewicz, L., Gal, E., Nyerges, E.A., Sablin, M. V., Brehard, S., Mashkour, M., Balasescu, A., Gillet, B., Hughes, S., Chassaing, O., Hitte, C., Vigne, J.-D., Dobney, K., Hanni, C., Bradley, D.G., Larson, G., 2016. Genomic and archaeological evidence suggests a dual origin of domestic dogs. Science. 352, 1228–1231.

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E. a., Marques-Bonet, T., Wayne, R.K., Novembre, J., 2014. Genome Sequencing Highlights the Dynamic Early History of Dogs. PLoS Genetics. 10, e1004016.

Galloy, J.M., 2010. PART V : WOODLAND PERIOD The Janey B . Goode Site (11S1232): Highlights of Investigations at a Massive Late Prehistoric Site in the American Bottom. Illinois Archaeology. 22, 529–552.

Gilbert, M.T.P., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., Carlson, J.E., Leebens-Mack, J.H., Schuster, S.C., 2007. Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. Nucleic Acids Research. 35, 1–10.

Giles, R.E., Blanc, H., Cann, H.M., Wallace, D.C., 1980. Maternal inheritance of human mitochondrial DNA. Proceedings of the National Academy of Sciences of the United States of America. 77, 6715–9.

Gill, K.M., Erlandson, J.M., 2014. The Island Chumash and Exchange in the Santa Barbara Channel Region. American Antiquity. 79, 570–572.

Gómez-Carballa, A., Olivieri, A., Behar, D.M., Achilli, A., Torroni, A., Salas, A., 2012. Genetic continuity in the Franco-Cantabrian region: New clues from autochthonous mitogenomes. PLoS ONE. 7, e32851.

Greig, K., Boocock, J., Prost, S., Horsburgh, K.A., Jacomb, C., Walter, R., Matisoo-Smith, E., 2015. Complete Mitochondrial Genomes of New Zealand's First Dogs. Plos ONE. 10, e0138536.

Guiry, E.J., 2012. Dogs as Analogs in Stable Isotope-Based Human Paleodietary Reconstructions: A Review and Considerations for Future Use. Journal of Archaeological Method and Theory. 19, 351–376.

Hagelberg, E., Clegg, J.B., 1991. Isolation and characterization of DNA from archaeological bone. Proceedings of the Royal Society B: Biological Sciences. 244, 45–50.

Hoffecker, J., Powers, W., Goebel, T., 1993. The colonization of Beringia and the peopling of the New World. Science. 259, 46–53.

Imes, D.L., Wictum, E.J., Allard, M.W., Sacks, B.N., 2012. Identification of single nucleotide polymorphisms within the mtDNA genome of the domestic dog to discriminate individuals with common HVI haplotypes. Forensic Science International Genetics. 6, 630–9.

Ishiguro, N., Nakano, M., 1996. Intra- and interbreed genetic variations of mitochondrial DNA major non- coding regions in Japanese native. Animal Genetics. 27, 397–405.

Jenkins, D.L., Davis, L.G., Stafford, T.W., Campos, P.F., Hockett, B., Jones, G.T., Cummings, L.S., Yost, C., Connolly, T.J., Yohe, R.M., Gibbons, S.C., Raghavan, M., Rasmussen, M., Paijmans, J.L.A., Hofreiter, M., Kemp, B.M., Barta, J.L., Monroe, C., Gilbert, M.T.P., Willerslev, E., 2012. Clovis Age Western Stemmed Projectile Points and Human Coprolites at the Paisley Caves. Science. 337, 223–228.

Jones, E.P., Eager, H.M., Gabriel, S.I., Johannesdottir, F., Searle, J.B., 2013. Genetic tracking of mice and other bioproxies to infer human history. Trends in Genetics. 29, 298–307.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. MapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics. 29, 1682–1684.

Kane, A.E., Robinson, C.K., 1988. Dolores Archaeological Program: Anasazi Communities at Dolores: McPhee Village. USDI Bureau of Reclamation, Engineering and Research Center, Denver.

Karlsson, E.K., Baranowska, I., Wade, C.M., Salmon Hillbertz, N.H.C., Zody, M.C., Anderson, N., Biagi, T.M., Patterson, N., Pielberg, G.R., Kulbokas, E.J., Comstock, K.E., Keller, E.T., Mesirov, J.P., von Euler, H., Kämpe, O., Hedhammar, A., Lander, E.S., Andersson, G., Andersson, L., Lindblad-Toh, K., 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. Nature Genetics. 39, 1321–8.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research. 30, 3059–3066.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution. 30, 772–780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28, 1647–1649.

Kelly, J.E., 1990. The Emergence of Mississippian Culture in the American Bottom Region. In: The Mississippian Emergence. pp. 113–152.

Kemp, B.M., Glenn Smith, D., 2010. Ancient DNA methodology: thoughts from Brian M. Kemp and David Glenn Smith on “Mitochondrial DNA of protohistoric remains of an Arikara population from South Dakota”. Human Biology. 82, 227–38.

Kemp, B.M., Schurr, T.G., 2010. Ancient and modern genetic variation in the Americas. In: Auerbach, B. (Ed.), Human Variation in the Americas: The Integration of Archeology and Biological Anthropology. Carbondale, IL, pp. 12–50.

Kemp, B.M., Smith, D.G., 2005. Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. Forensic Science International. 154, 53–61.

Kim, K.S., Lee, S.E., Jeong, H.W., Ha, J.H., 1998. The complete nucleotide sequence of the domestic dog (Canis familiaris) mitochondrial genome. Molecular Phylogenetics and Evolution. 10, 210–20.

Kitchen, A., Miyamoto, M.M., Mulligan, C.J., 2008. A three-stage colonization model for the peopling of the Americas. PloS ONE. 3, e1596.

Kuckelman, K.A., 2003. The Archaeology of Yellow Jacket Pueblo (Site 5MT5): Excavations at a Large Community Center in Southwestern Colorado [WWW Document]. http://www.crowcanyon.org/yellowjacket.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution. 29, 1695–1701.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods. 9, 357–359.

Lapham, H.A., 2010. A Baumer Phase Dog Burial from the Kincaid Site in Southern Illinois. Illinois Archaeology. 22, 437–463.

Larson, G., Cucchi, T., Fujita, M., Matisoo-Smith, E., Robins, J., Anderson, A., Rolett, B., Spriggs, M., Dolman, G., Kim, T.-H., Thuy, N.T.D., Randi, E., Doherty, M., Due, R.A., Bollt, R., Djubiantono, T., Griffin, B., Intoh, M., Keane, E., Kirch, P., Li, K.-T., Morwood, M., Pedriña, L.M., Piper, P.J., Rabett, R.J., Shooter, P., Van den Bergh, G., West, E., Wickler, S., Yuan, J., Cooper, A., Dobney, K., 2007. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proceedings of the National Academy of Sciences of the United States of America. 104, 4834–9.

Larson, G., Karlsson, E.K., Perri, A., Webster, M.T., Ho, S.Y.W., Peters, J., Stahl, P.W., Piper, P.J., Lingaas, F., Fredholm, M., Comstock, K.E., Modiano, J.F., Schelling, C., Agoulnik, A.I., Leegwater, P. a, Dobney, K., Vigne, J.-D., Vilà, C., Andersson, L., Lindblad-Toh, K., 2012. Rethinking dog domestication by integrating genetics, archeology, and biogeography. Proceedings of the National Academy of Sciences of the United States of America. 109, 8878–83.

Lee, E.J., Merriwether, D.A., Kasparov, A.K., Nikolskiy, P.A., Sotnikova, M. V., Yu Pavlova, E., Pitulko, V. V.,

2015. Ancient DNA analysis of the oldest Canid species from the Siberian Arctic and genetic contribution to the domestic dog. PLoS ONE. 10, 1–13.

Leonard, J. A., Wayne, R.K., Wheeler, J., Valadez, R., Guillén, S., Vilà, C., 2002. Ancient DNA evidence for Old World origin of New World dogs. Science. 298, 1613–6.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25, 2078–2079.

Li, Q., Liu, Z., Li, Y., Zhao, X., Dong, L., Pan, Z., Sun, Y., Li, N., Xu, Y., Xie, Z., 2008. Origin and phylogenetic analysis of Tibetan Mastiff based on the mitochondrial DNA sequence. Journal of Genetics and Genomics. 35, 335–40.

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E. A., Ponting, C.P., Galibert, F., Smith, D.R., DeJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.-W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.-P., Parker, H.G., Pollinger, J.P., Searle, S.M.J., Sutter, N.B., Thomas, R., Webber, C., Baldwin, J., Abebe, A., Abouelleil, A., Aftuck, L., Ait-Zahra, M., Aldredge, T., Allen, N., An, P., Anderson, S., Antoine, C., Arachchi, H., Aslam, A., Ayotte, L., Bachantsang, P., Barry, A., Bayul, T., Benamara, M., Berlin, A., Bessette, D., Blitshteyn, B., Bloom, T., Blye, J., Boguslavskiy, L., Bonnet, C., Boukhgalter, B., Brown, A., Cahill, P., Calixte, N., Camarata, J., Cheshatsang, Y., Chu, J., Citroen, M., Collymore, A., Cooke, P., Dawoe, T., Daza, R., Decktor, K., DeGray, S., Dhargay, N., Dooley, K., Dooley, K., Dorje, P., Dorjee, K., Dorris, L., Duffey, N., Dupes, A., Egbiremolen, O., Elong, R., Falk, J., Farina, A., Faro, S., Ferguson, D., Ferreira, P., Fisher, S., FitzGerald, M., Foley, K., Foley, C., Franke, A., Friedrich, D., Gage, D., Garber, M., Gearin, G., Giannoukos, G., Goode, T., Goyette, A., Graham, J., Grandbois, E., Gyaltsen, K., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Healy, C., Hegarty, R., Honan, T., Horn, A., Houde, N., Hughes, L., Hunnicutt, L., Husby, M., Jester, B., Jones, C., Kamat, A., Kanga, B., Kells, C., Khazanovich, D., Kieu, A.C., Kisner, P., Kumar, M., Lance, K., Landers, T., Lara, M., Lee, W., Leger, J.-P., Lennon, N., Leuper, L., LeVine, S., Liu, J., Liu, X., Lokyitsang, Y., Lokyitsang, T., Lui, A., Macdonald, J., Major, J., Marabella, R., Maru, K., Matthews, C., McDonough, S., Mehta, T., Meldrim, J., Melnikov, A., Meneus, L., Mihalev, A., Mihova, T., Miller, K., Mittelman, R., Mlenga, V., Mulrain, L., Munson, G., Navidi, A., Naylor, J., Nguyen, T., Nguyen, N., Nguyen, C., Nguyen, T., Nicol, R., Norbu, N., Norbu, C., Novod, N., Nyima, T., Olandt, P., O'Neill, B., O'Neill, K., Osman, S., Oyono, L., Patti, C., Perrin, D., Phunkhang, P., Pierre, F., Priest, M., Rachupka, A., Raghuraman, S., Rameau, R., Ray, V., Raymond, C., Rege, F., Rise, C., Rogers, J., Rogov, P., Sahalie, J., Settipalli, S., Sharpe, T., Shea, T., Sheehan, M., Sherpa, N., Shi, J., Shih, D., Sloan, J., Smith, C., Sparrow, T., Stalker, J., Stange-Thomann, N., Stavropoulos, S., Stone, C., Stone, S., Sykes, S., Tchuinga, P., Tenzing, P., Tesfaye, S., Thoulutsang, D., Thoulutsang, Y., Topham, K., Topping, I., Tsamla, T., Vassiliev, H., Venkataraman, V., Vo, A., Wangchuk, T., Wangdi, T., Weiand, M., Wilkinson, J., Wilson, A., Yadav, S., Yang, S., Yang, X., Young, G., Yu, Q., Zainoun, J., Zembek, L., Zimmer, A., Lander, E.S., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature. 438, 803–19.

Lindgreen, S., 2012. AdapterRemoval: Easy Cleaning of Next Generation Sequencing Reads. BMC Research Notes. 5, 337.

Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S.,

Valdiosera, C., Richards, S.M., Rohlach, A., Barreto-Romero, M.I., Flores-Espinoza, I., Tomasto-Cagigao, E., Watson-Jimenez, L., Makowski, K., Santiago-Leboreiro-Reyna, I., Mansilla-Lory, J., Ballivian-Torrez, J.A., Rivera, M.A., Burger, R.L., Constanza-Ceruti, M., Reinhard, J., Wells, R.S., Politis, G., Santoro, C.M., Standen, V.G., Smith, C., Reich, D., Ho, S.Y.W., Cooper, A., Haak, W., 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Science Advances. 2, e1501385.

Loog, L., Thalmann, O., Sinding, M.-H.S., Shuenemann, V.J., Perri, A., Germonpre, M., Bocherens, H., Witt, K.E., Samaniego Castruita, J.A., Velasco, M.S., Lundstroem, I.K.C., Wales, N., Sonet, G., Frantz, L., Schroeder, H., Budd, J., Jimenez, E., Federov, S., Gasparyan, B., Kandel, A.W., Laznickova-Galetova, M., Napierala, H., Uerpmann, H.-P., Nikolskiy, P.A., Pavlova, E.Y., Pitulko, V. V., Herzig, K.-H., Malhi, R.S., Willerslev, E., Hansen, A.J., Dobney, K., Gilbert, M.T.P., Krause, J., Larson, G., Eriksson, A., Manica, A., 2017. Modern wolves trace their ancestry to a late Pleistocene expansion from Beringia. In Review.

Malhi, R.S., Eshleman, J. A., Greenberg, J. A., Weiss, D. A., Schultz Shook, B. A., Kaestle, F. A., Lorenz, J.G., Kemp, B.M., Johnson, J.R., Smith, D.G., 2002. The structure of diversity within New World mitochondrial DNA haplogroups: implications for the prehistory of North America. American Journal of Human Genetics. 70, 905–19.

Malyarchuk, B., Derenko, M., Denisova, G., Maksimov, A., Wozniak, M., Grzybowski, T., Dambueva, I., Zakharov, I., 2011. Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. Journal of Human Genetics. 56, 583–588.

Masson, M.A., Lope, C.P., 2008. Animal use at the Postclassic Maya center of Mayapan. Quaternary International. 191, 170–183.

Matisoo-Smith, E., Robins, J.H., 2004. Origins and dispersals of Pacific peoples: evidence from mtDNA phylogenies of the Pacific rat. Proceedings of the National Academy of Sciences of the United States of America. 101, 9167–72.

Meltzer, D.J., Holliday, V.T., 2010. Would North American Paleoindians have Noticed Younger Dryas Age Climate Changes? Journal of World Prehistory. 23, 1-41.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop, GCE 2010. 1–8.

Moore, C.B., 1915. Aboriginal sites on the Tennessee River. Journal of the Academy for Natural Sciences of Philadelphia. 16, 169–428.

Morey, D.F., 2006. Burying key evidence : the social bond between dogs and people. Journal of Archaeological Science. 33, 158–175.

Morey, D.F., Wiant, M.D., 1992. Early Holocene Domestic Dog Burials from the North American Midwest. Current Anthropology. 33, 224–230.

Mulligan, C.J., Kitchen, A., Miyamoto, M.M., 2008. Updated three-stage model for the peopling of the Americas. PloS ONE. 3, e3199.

Noe-Nygaard, N., 1988. δ 13 C-values of dog bones reveal the nature of changes in man's food resources at the mesolithic-neolithic transition, Denmark. Chemical Geology: Isotope Geoscience section. 13, 87–96.

Overstreet, D.F., Kolb, M.F., 2003. Geoarchaeological contexts for Late Pleistocene Archaeological Sites with Human-modified Woolly Mammoth Remains in Southeastern Wisconsin. Geoarchaeology. 18, 91–114.

Ovodov, N.D., Crockford, S.J., Kuzmin, Y. V, Higham, T.F.G., Hodgins, G.W.L., van der Plicht, J., 2011. A 33,000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the Last Glacial Maximum. PloS ONE. 6, e22821.

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A., Luo, L.-Y., Angleby, H., Ardalan, A., Ekström, C., Sköllermo, A., Lundeberg, J., Matsumura, S., Leitner, T., Zhang, Y.-P., Savolainen, P., 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. Molecular Biology and Evolution. 26, 2849–64.

Parker, H.G., Dreger, D.L., Rimbault, M., Davis, B.W., Mullen, A.B., Carpintero-Ramirez, G., Ostrander, E.A., 2017. Genomic Analyses Reveal the Influence of Geographic Origin, Migration, and Hybridization on Modern Dog Breed Development. Cell Reports. 19, 697–708.

Parmalee, P.W., 1959. Appendix II: animals remains from the Modoc Rock Shelter site, Randolph County, Illinois, Summary Report of Modoc Rock Shelter 1952, 1953, 1955, 1956, Report of Investigations 8. Springfield, Illinois.

Parmalee, P.W., Paloumpis, A.A., Wilson, N., 1972. Animals Utilized by Woodland Peoples Occupying the Apple Creek Site, Illinois, Research Papers 5. Springfield, I.

Pauketat, T.R., 2004. Ancient Cahokia and the Mississippians. Cambridge University Press.

Pauketat, T.R., Alt, S.M., 2015. Medieval Mississippians: The Cahokian World.

Pauketat, T.R., Emerson, T.E., 1997. Cahokia: Domination and Ideology in the Mississippian World. University of Nebraska Press.

Perego, U. A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., Hooshiar Kashani, B., Ritchie, K.H., Scozzari, R., Kong, Q.-P., Myres, N.M., Salas, A., Semino, O., Bandelt, H.-J., Woodward, S.R., Torroni, A., 2009. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Current biology. 19, 1–8.

Perry, J., 2012. Diversity, Exchange, and Complexity in the California Bight. The Oxford Handbook of North American Archaeology.

Poinar, H.N., 2003. The top 10 list: criteria of authenticity for DNA from ancient and forensic samples. International Congress Series. 1239, 575–579.

Potter, M.A., Baby, R.S., 1964. Hopewellian Dogs. The Ohio Journal of Science. 64, 36–40.

Raff, J.A., Bolnick, D.A., Tackney, J., O'Rourke, D.H., 2011. Ancient DNA perspectives on American

colonization and population history. American Journal of Physical Anthropology. 146, 503–514.

Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir, V., Yadav, R., Malaspinas, A.-S., White, S.S., Allentoft, M.E., Cornejo, O.E., Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M., Korneliussen, T.S., Meltzer, D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S., Brunak, S., Sicheritz-Ponten, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica, A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., Willerslev, E., 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature. 506, 225–9.

Reich, D., Patterson, N., Campbell, D., 2012. Reconstructing native American population history. Nature. 488, 370-374.

Reynolds, A.W., Raff, J.A., Bolnick, D.A., Cook, D.C., Kaestle, F.A., 2015. Ancient DNA from the Schild site in Illinois: Implications for the Mississippian transition in the lower Illinois River Valley. American Journal of Physical Anthropology. 156, 434–448.

Rick, T.C., Culleton, B.J., Smith, C.B., Johnson, J.R., Kennett, D.J., 2011. Stable isotope analysis of dog, fox, and human diets at a Late Holocene Chumash village (CA-SRI-2) on Santa Rosa Island, California. Journal of Archaeological Science. 38, 1385–1393.

Rick, T.C., Erlandson, J.M., Vellanoweth, R.L., Braje, T.J., 2005. From pleistocene mariners to complex hunter-gatherers: The archaeology of the california channel islands. Journal of World Prehistory. 19, 169–228.

Rick, T.C., Walker, P.L., Willis, L.M., Noah, A.C., Erlandson, J.M., Vellanoweth, R.L., Braje, T.J., Kennett, D.J., 2008. Dogs, humans and island ecosystems: the distribution, antiquity and ecology of domestic dogs (Canis familiaris) on California's Channel Islands, USA. The Holocene. 18, 1077–1087.

Sacks, B.N., Brown, S.K., Stephens, D., Pedersen, N.C., Wu, J.-T., Berry, O., 2013. Y chromosome analysis of dingoes and southeast asian village dogs suggests a neolithic continental expansion from southeast Asia followed by multiple austronesian dispersals. Molecular Biology and Evolution. 30, 1103–18.

Schurr, T.G., Sherry, S.T., 2004. Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. American Journal of Human Biology. 16, 420–39.

Schwartz, M., 2000. The Form and Meaning of Maya and Mississippian Dog Representations. In: Dogs Through Time: An Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. pp. 217–226.

Seeman, M.F., 1979. Hopewell Interaction Sphere: the Evidence for Interregional Trade and Structural Complexity.

Shahid, S.A., Xiao, Y., Khan, S., Feng, D., Johnson, G.S., Ha, J., 2004. Sequence Diversity of the Canine Mitochondrial Genome. Unpublished.

Shannon, L.M., Boyko, R.H., Castelhano, M., Corey, E., Hayward, J.J., Mclean, C., White, M.E., Abi, M., Anita, B.A., Ikombe, N., Calero, J., Galov, A., 2015. Genetic structure in village dogs reveals a Central Asian domestication origin. Proceedings of the National Academy of Sciences of the United States of America. 112, 13639-644

Slater, P. A., Hedman, K.M., Emerson, T.E., 2014. Immigrants at the Mississippian polity of Cahokia: strontium isotope evidence for population movement. Journal of Archaeological Science. 44, 117–127.

Smith, B.D., 2007. The Mississippian Emergence. University of Alabama Press.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30, 1312–1313.

Stewart, F.L., Stewart, K.M., 1996. The Boardwalk and Grassy Bay Sites: Patterns of Seasonality and Subsistance on the Northern Northwest Coast, B.C. Canadian Journal of Archaeology. 20, 39–60.

Stoneking, M., 2000. Hypervariable sites in the mtDNA control region are mutational hotspots. American Journal of Human Genetics. 67, 1029–1032.

Struever, S., Houart, G.L., 1972. Analysis of the Hopewell interaction sphere.

Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D.G., Mulligan, C.J., Bravi, C.M., Rickards, O., Martinez-Labarga, C., Khusnutdinova, E.K., Fedorova, S. A., Golubenko, M. V, Stepanov, V. A., Gubina, M. A., Zhadanov, S.I., Ossipova, L.P., Damba, L., Voevoda, M.I., Dipierri, J.E., Villems, R., Malhi, R.S., 2007. Beringian standstill and spread of Native American founders. PloS one. 2, e829.

Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpre, M.., Sablin, M.V., Lopez-Giraldez, F., Domingo-Roura, X., Napierala, H., Uerpmann, H.-P., Loponte, D.M., Acosta, A.., Giemsch, L., Schmitz, R.W., Worthington, B., Buikstra, J.E., Druzhkova, A., Graphodatsky, A.S., Ovodov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfli, K.-P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., Wayne, R.K., 2013. Complete Mitochondrial Genomes of Ancient Canids Suggest a European Origin of Domestic Dogs. Science. 342, 871–874.

Thomson, V.A., Lebrasseur, O., Austin, J.J., Hunt, T.L., Burney, D.A., Denham, T., Rawlence, N.J., Wood, J.R., Gongora, J., Girdland Flink, L., Linderholm, A., Dobney, K., Larson, G., Cooper, A., 2014. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proceedings of the National Academy of Sciences. 111, 4826–4831.

Torres, M.M., Bravi, C.M., Bortolini, M.-C., Duque, C., Callegari-Jacques, S., Ortiz, D., Bedoya, G., Groot de Restrepo, H., Ruiz-Linares, A., 2006. A revertant of the major founder Native American haplogroup C common in populations from northern South America. American Journal of Human Biology. 18, 59–65.

Torroni, a, Schurr, T.G., Cabell, M.F., Brown, M.D., Neel, J. V, Larsen, M., Smith, D.G., Vullo, C.M., Wallace, D.C., 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. American Journal of Human Genetics. 53, 563–90.

van Asch, B., Zhang, A.-B., Oskarsson, M.C.R., Klütsch, C.F.C., Amorim, A., Savolainen, P., 2013. Pre-Columbian origins of Native American dog breeds, with only limited replacement by European dogs, confirmed by mtDNA analysis. Proceedings of the Royal Society B: Biological Sciences. 280, 20131142.

Verdu, P., Pemberton, T.J., Laurent, R., Kemp, B.M., Gonzalez-Oliver, A., Gorodezky, C., Hughes, C.E., Shattuck, M.R., Petzelt, B., Mitchell, J., Harry, H., William, T., Worl, R., Cybulski, J.S., Rosenberg, N.A., Malhi, R.S., 2014. Patterns of admixture and population structure in native populations of Northwest North America. PLoS Genetics. 10, e1004530.

Verginelli, F., Capelli, C., Coia, V., Musiani, M., Falchetti, M., Ottini, L., Palmirotta, R., Tagliacozzo, A., De Grossi Mazzorin, I., Mariani-Costantini, R., 2005. Mitochondrial DNA from prehistoric canids highlights relationships between dogs and South-East European wolves. Molecular Biology and Evolution. 22, 2541–51.

Verscheure, S., Backeljau, T., Desmyter, S., 2014. Dog mitochondrial genome sequencing to enhance dog mtDNA discrimination power in forensic casework. Forensic Science International: Genetics. 12, 60-68.

Vonholdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E., Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D. A., Auton, A., Reynolds, A., Bryc, K., Brisbin, A., Knowles, J.C., Mosher, D.S., Spady, T.C., Elkahloun, A., Geffen, E., Pilot, M., Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.-L., Zhang, Y.-P., Bustamante, C.D., Ostrander, E. a, Novembre, J., Wayne, R.K., 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature. 464, 898–902.

Wang, S., Lewis, C.M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M. V, Molina, J. A., Gallo, C., Mazzotti, G., Poletti, G., Hill, K., Hurtado, A.M., Labuda, D., Klitz, W., Barrantes, R., Bortolini, M.C., Salzano, F.M., Petzl-Erler, M.L., Tsuneto, L.T., Llop, E., Rothhammer, F., Excoffier, L., Feldman, M.W., Rosenberg, N. a, Ruiz-Linares, A., 2007. Genetic variation and population structure in native Americans. PLoS Genetics. 3, e185.

Webb, K.M., Allard, M.W., 2009. Mitochondrial genome DNA analysis of the domestic dog: identifying informative SNPs outside of the control region. Journal of Forensic Sciences. 54, 275–88.

Webb, W.S., DeJarnette, D.L., 1942. An Archaeological Survey of Pickwick Basin in the Adjacent Portions of the States of Alabama, Mississippi, and Tennessee, Bulletin 129. Washington, D.C.

Webb, W.S., DeJarnette, D.L., 1948a. The Flint River Site, Museum Paper 23. Tuscaloosa, Alabama.

Webb, W.S., DeJarnette, D.L., 1948b. Little Bear Creek, Museum Paper 26. Tuscaloosa, Alabama.

Wei, Z., Wang, W., Hu, P., Lyon, G.J., Hakonarson, H., 2011. SNVer: A statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. Nucleic Acids Research. 39, 1–13.

White, C.D., Pohl, M.E.D., Schwarcz, H.P., Longstaffe, F.J., 2001. Isotopic Evidence for Maya Patterns of Deer and Dog Use at Preclassic Colha. Journal of Archaeological Science. 28, 89–107.

Willerslev, E., Cooper, A., 2005. Ancient DNA. Proceedings of the Royal Society B: Biological Sciences. 272, 3–16.

Witt, K.E., Judd, K., Kitchen, A., Grier, C., Kohler, T. A., Ortman, S.G., Kemp, B.M., Malhi, R.S., 2015. DNA analysis of ancient dogs of the Americas: Identifying possible founding haplotypes and reconstructing population histories. Journal of Human Evolution. 79, 105–118.

Zegura, S.L., Karafet, T.M., Zhivotovsky, L. A., Hammer, M.F., 2004. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Molecular Biology and Evolution. 21, 164–75.

CHAPTER FOUR: ASSESSING DIET IN LATE WOODLAND AND MISSISSIPPIAN DOGS IN THE AMERICAN BOTTOM THROUGH ISOTOPIC ANALYSIS AND DNA SEQUENCING

Abstract

Maize agriculture was an important component of Mississippian society. Maize is thought to have arrived to the American Bottom (the Mississippi River Valley in Southern Illinois) around 900 AD, and was grown and consumed intensively during the Mississippian period. However, the exact timing of maize arrival, as well as when it became a primary part of the Mississippian diet, is unknown. Dog remains recovered from Janey B. Goode, a site near Cahokia, dating to the Late Woodland-Mississippian transition, can be useful as a proxy for understanding human diet during this transition. We analyzed the dietary isotopes ^{13}C, ^{15}N, and ^{18}O in dog teeth and bones to assess how the population's diet changed over time, and also sequenced DNA from dog coprolites from the site to identify specific dietary components. The dogs were found to have a diet of mostly C_3 plants during the Late Woodland period, with a diet high in C_4 plants during the Mississippian period, starting around 1000 AD. DNA sequences that mapped to maize were also identified in 37.5% of the coprolites, along with DNA sequences of a number of other native species, including crops like squash, nightshade, and tobacco, and fish including sunfish and gizzard shad. This study shows the utility of dogs as indicators of human diet, and suggests that humans in the American Bottom likely ate large amounts of maize starting at the beginning of the Mississippian period, and supplemented that with other crops and Mississippi River fish.

Introduction

The shift from the Late Woodland to Mississippian periods in the American Bottom region of Illinois (Figure 4.1) is marked by a number of lifestyle changes in the people who occupied this region, including an increase in population density and village size, a shift from hunting and gathering with small-scale horticulture to maize agriculture, and the rise of Cahokia, a large mound city with far-reaching cultural influences (Pauketat, 2004). During the Mississippian period, a number of more rural communities were abandoned as people moved closer to Cahokia, and communities became more centralized, with small hamlets all connected to a single mound, which functioned as a community center (Emerson, 1997). Additionally, the number of trade interactions between populations increased and similar iconography and religious motifs were found across the eastern half of the United States (Kelly, 2007).

Maize likely arrived to the region around 1100 ybp, but intensive maize agriculture did not occur until 1000-950 ybp (Buikstra and Milner, 1989; Kidder, 1992; Hart, 1999; Cook and Schurr, 2009; Vanderwarker et al., 2013). Prior to the arrival of maize, a number of local crops had already been domesticated, which were collectively named the Eastern Agricultural Complex: goosefoot (*Chenopodium berlandieri*), maygrass (*Phalaris caroliniana*), little barley (*Hordeum pusillum*), and sumpweed (*Iva annua*) (Smith, 1989; Lopinot, 1997; Simon and Lopinot, 2006; Vanderwarker et al., 2013; VanDerwarker et al., 2016). A number of other crops were grown in the American Bottom during the Late Woodland period, including squash (*Cucurbita pepo*), sunflower (*Helianthus anuus*), and tobacco (*Nicotiana spp.*) and nearly all of these continued to be cultivated during the Mississippian period, with maize acting as a complementary crop, rather than a replacement crop (Lopinot, 1997; Simon and Lopinot, 2006). The archaeological sites with the earliest identified maize in Eastern North America are the Icehouse Bottom site in Tennessee (1700 years before present(ybp)), the Edwin Harness site in Ohio (1700 ybp), and the Grand Banks site in Ontario (1500 ybp) (Crawford et al., 1997). It has also been suggested that maize had been present in the American Bottom since the Middle Woodland period (as early as 1500 ybp) (Riley et

al., 1994), but these samples have been demonstrated to either be younger than originally thought, or a crop plant other than maize (Simon, 2017). Maize consumption varied at different Mississippian sites, but seems to have peaked during the Late Mississippian period (Buikstra and Milner, 1989; Lopinot, 1997; Hedman et al., 2002; Simon and Lopinot, 2006). At Cahokia, there were differences in diet that reflected social stratification (Ambrose et al., 2003), but maize was consumed heavily even at Fort Ancient and Oneonta sites, which were often considered to be politically and culturally separate from Cahokia (Buikstra and Milner, 1989; Emerson et al., 2005; Cook and Price, 2015).

Given this connection between maize agriculture and the rise of the Mississippians, determining the timing of increased maize consumption is important for understanding the Mississippian transition (Schroeder, 1999). It has long been debated whether the introduction of maize was the catalyst for the development of the stratified society at Cahokia, or if maize was found to be a useful crop for feeding large populations after Cahokia was already established (Kidder, 1992; Lopinot, 1997). The timing of increased maize consumption in humans could help resolve this question – if maize consumption did not increase until after the founding of Cahokia (after 1000 ybp), the latter is more likely. However, if maize consumption intensified during the Terminal Late Woodland period (before 1000 ybp), it may have played a pivotal role in the development of Cahokia and the Mississippian culture. Given that maize likely arrived to the American Bottom around 1100 ybp (Vanderwarker et al., 2013; Simon, 2017), examining the diet of humans who lived during the Terminal Late Woodland and Early Mississippian periods would provide insight as to when heavy maize consumption began. However, most archaeological sites that span the Late Woodland to Mississippian periods only have human burials that postdate the transition to maize agriculture; human burial practice was uncommon during the Late Woodland period (Fortier and Jackson, 2000). The domestic dog was an important part of human communities during the Late Woodland period, and dogs were buried in large numbers at multiple archaeological sites in the American Bottom (Parmalee et al., 1972; Cantwell, 1980; Borgic and Galloy, 2004; Lapham, 2010). Given the close relationship between

dogs and humans, these dogs could be used to indirectly examine the diet of the humans from the American Bottom.

Various methods can be used to assess diet in an ancient population. Remnants of seeds, bones, and pollen can be analyzed to identify which species were present at a site. Stable isotope analysis of carbon, nitrogen, and oxygen isotopes in bone collagen and apatite can be used to distinguish between different diets. The amount of ^{13}C a plant contains is dependent on its mode of photosynthesis, categorizing it as either a C_3 or C_4 plant. Plants using the dicarboxylic acid pathway of photosynthesis (C_4) are more enriched in ^{13}C than plants that use the Calvin pathway of photosynthesis (C_3) (Park and Epstein, 1960; van der Merwe and Vogel, 1978). C_3 plants such as forest plants and the crops utilized by Middle and Late Woodland period populations (including sumpweed, goosefoot and squash) will have a smaller $\delta^{13}C$ (averaging -26.5‰) than C_4 plants such as warm season grasses, including maize (ranging from -14‰ to -9‰) (Ambrose, 1990). There are few C_4 plants in the American Bottom other than maize, which makes ^{13}C a useful isotope for differentiating between the crops being consumed in the Middle and Late Woodland periods and maize (van der Merwe and Vogel, 1978). The amount of ^{15}N in an animal or plant is determined by its trophic level; plants have less $\delta^{15}N$ than herbivores, which have less $\delta^{15}N$ than carnivores (Schoeninger and DeNiro, 1984). A shift from hunting and gathering to maize agriculture will be marked by an increase in $\delta^{13}C$. Humans from the Archaic period through the end of the Woodland period (9000-1000 ybp) had average $\delta^{13}C$ values of -20‰, indicating a diet of C_3 plants (van der Merwe and Vogel, 1978), while humans in the Mississippian period show increasing $\delta^{13}C$ over time, with values of -18 to -15‰ during the Early Mississippian period (1000-800 ybp), and -9‰ during the Late Mississippian period (600-460 ybp) (Buikstra and Milner, 1989; Hedman et al., 2002; Schoeninger, 2009; Yerkes, 2011). In some Mississippian sites, some humans were identified as eating a diet of up to 80% C_4 plants (like maize) (Hedman et al., 2002; Ambrose et al., 2003), showing an increase in plant consumption over time. The fraction of ^{18}O in an organism is a reflection of the water they consumed, and can be a

measure of climate and seasonality, as well as diet (Sponheimer and Lee-Thorp, 1999; Balasse et al., 2002, 2003; White et al., 2004). Water contained within plants undergoes evapotranspiration, which enriches the water that remains with ^{18}O (Sponheimer and Lee-Thorp, 1999; White et al., 2004). Organisms that eat more plants often drink less ground water, so their $\delta^{18}O$ values will be higher than those that drink more water (Bocherens et al., 1996; White et al., 2004; Bowen et al., 2009). Therefore, the $\delta^{18}O$ value of an organism can be used to estimate protein consumption, with a lower $\delta^{18}O$ value indicating greater animal protein consumption.

Collagen from bone and hydroxyapatite from bone and tooth enamel can be useful in assessing diet, and bone collagen, bone apatite, and enamel apatite all reflect different aspects of an organism's diet. Collagen contains ^{13}C and ^{15}N, while apatite contains ^{13}C and ^{18}O. The diet is routed differently into different tissues – the collagen isotope values reflect the protein component of the diet, while the apatite isotope values reflect the whole diet (Ambrose and Norr, 1993; Tykot et al., 1996; Ambrose and Katzenberg, 2000). Bone and enamel apatite reflect diet from different periods of an individual's life – enamel apatite reflects the diet from when the teeth were forming (ie as juveniles), while bone apatite reflects the diet as an adult (Tykot et al., 1996). Typically, bone and enamel $\delta^{18}O$ values should be identical, but because the hydroxyapatite in bone is more prone to diagenesis, the enamel $\delta^{18}O$ value is considered to be a more accurate measure (Bocherens et al., 1996; Wright and Schwarcz, 1999).

Bone apatite and bone collagen have different $\delta^{13}C$ offsets relative to the $\delta^{13}C$ of the diet, likely because they reflect diet at different developmental ages (post- and pre-weaning). The offset of bone apatite from total diet in controlled diet experiments with rats is 9.4‰ (Ambrose and Norr, 1993), while the offset of bone collagen from diet in a diet in which protein and non-protein macronutrients have the same $\delta^{13}C$ is 5‰. If the difference between $\delta^{13}C$ in collagen and apatite is greater than 4‰, it indicates that the dietary composition of the whole diet differs from the dietary composition of the protein source. For example, if the dogs were eating both maize and woodland animals, the animals they ate would have

had likely eaten C_3 plants, so the dogs' collagen $\delta^{13}C$ value would be lower than the apatite $\delta^{13}C$ value.

Dog bones have often been analyzed for dietary isotopes to infer human diet. One of the earliest examples of comparing dog and human isotopes was a study which compared hunter-gatherer and farmer diets in Denmark, and found that dogs and humans from the same archaeological site had similar $\delta^{13}C$ values (Noe-Nygaard, 1988). Dogs and humans from the same archaeological site have been demonstrated to have similar diets, and therefore similar stable isotope values (Tankersley and Koster, 2009; Guiry, 2012). Generally, dogs have been shown to be a good proxy for $\delta^{13}C$, but slightly less so for $\delta^{15}N$ (Guiry, 2012) – this discrepancy is partially due to the fact that humans have a naturally high $\delta^{15}N$ compared to other animals with similar diets, although the cause of this offset is largely unknown (Fizet et al., 1995; Hedges and Reynard, 2007). Dogs have also been used as a proxy for humans to specifically assess maize consumption in multiple regions of North America. In Josey Farm, a Mississippian site from Mississippi, stable isotopes of the dogs were used to demonstrate that humans from Josey Farm likely ate maize, and had similar diets to other nearby Mississippian sites (Hogue, 2003). At the Mayan city of Colha, dogs showed an increase in maize consumption over time, suggesting that they were provisioned by humans, unlike deer from the same site that consumed only C_3 plants (White et al., 2001). Also, analysis of dogs from two coastal Mississippian sites in New Jersey demonstrated that, although maize had not been previously identified in the area, humans at these sites likely supplemented their marine diet with maize (Allitt et al., 2009). Dogs show promise as a dietary proxy for humans to date the intensification of maize consumption in the American Bottom, because dog and human diets from the same archaeological site are often highly similar.

Another approach to assessing diet in a population is to sequence DNA from coprolites, or ancient feces. Unlike isotopic analysis, which assesses broad-scale trends in diet, DNA sequencing of coprolites can identify specific taxa down to the order, family, or even genus or species level. Coprolites only record recent meals, not diet over long periods of time, but can show specific dietary components, and so

coprolite analysis can complement stable isotope analysis. A study of cave sloth coprolites identified multiple orders of plants, including many taxa that are still found at the site they were recovered from (Hofreiter et al., 2000, 2003). Studies of other species including moa (Wood et al., 2008), cave hyena (Bon et al., 2012), and Balearic mountain goats (Welker et al., 2014), have identified DNA from both dietary components and the species that produced the coprolite, which has helped reconstruct ancient food webs. Maize DNA in coprolites can be used to determine whether dogs were eating maize, and the coprolites can be directly radiocarbon dated to estimate when maize arrived to the area.

The coprolites can also be used to identify other dietary components of the dogs, to infer what humans were eating in addition to maize. According to the archaeological record, humans were eating a varied diet of crop plants (such as sumpweed, squash, and maize), wild plants (such as pawpaw, hazelnuts, and wild grape), fish (including gar, perch, and catfish), waterfowl (including multiple species of duck), and terrestrial mammals (such as deer and squirrels) during the Mississippian period (Table 4.6) (Smith, 1989; Kelly, 1997; Lopinot, 1997; Simon and Lopinot, 2006; Simon, 2010; Yerkes, 2011; Vanderwarker et al., 2013). In general, marine vertebrates seem to have been consumed in greater numbers than terrestrial vertebrates (Kelly, 1997). However, finding plant remnants or animal bones at an archaeological site does not necessarily imply that the species was being eaten, unless there is evidence of cooking or cut marks on bones. By identifying taxa other than maize in the coprolites, it would be possible to confirm specific taxa as being part of the dog diet, and by proxy, the human diet as well.

The goal of this study is to use dogs from the American Bottom to test hypotheses about the timing of maize arrival. First, bone collagen, bone apatite, and enamel apatite from dog skeletal remains from the Late Woodland and Mississippian period were analyzed to determine when dogs started to consume maize. I found that dogs from Janey B. Goode begin to consume large amounts of maize around 990 ybp, close to the start of the Mississippian period. Second, high-throughput sequencing of dog coprolites was used to identify specific components of the dogs' diet, and to determine if maize can be

identified in the coprolites. Maize was identified in the coprolites, and many other species were identified including fish, internal parasites, and crops that were native to the American Bottom. The isotope and DNA sequencing data demonstrate an increase in maize consumption over time, and show that humans at Janey B. Goode likely had a varied diet in addition to the maize they ate during the Mississippian period.

Methods

Samples

The samples used in this study are from the Janey B. Goode site (11S1232), located in Brooklyn, IL, which is only eight kilometers from Cahokia (Galloy, 2010). Janey B. Goode was occupied from the Woodland through the Mississippian periods. Dozens of dogs were deliberately buried at this site during the Late Woodland, including some that were beheaded prior to burial, perhaps as part of a ritual (Galloy, 2010). Dog remains have also been recovered during the Mississippian period, but are more fragmentary, likely because dogs were being consumed during this time (Schwartz, 2000; Galloy, 2010). Bones (primarily ribs, but also a humerus and a radius) from twelve dogs were selected based on their provenience, with the hope that these dogs represented the full span of human occupation of the site. Eight teeth (all lower M3s) were also selected for these twelve individuals, when they were available. The list of individuals used in this study, along with their radiocarbon age (which was measured by the Illinois State Archaeological Survey), can be found in Table 4.1.

Approximately 100 dog coprolites have also been recovered from the Janey B. Goode site, which is unique among archaeological sites in eastern North America (Fortier, 2015). The coprolites have been identified as dog due to their contents, which include many bones and fish scales, and have white interiors, reflecting the consumption of large amounts of bone. A macroscopic analysis of some of the coprolites from the site has been completed, and the taxa that have been identified include rodents, birds, and fish, specifically gar, as well as unidentifiable plant fibers (Fortier, 2015). Ten of these coprolites, which were not previously radiocarbon dated, were selected for analysis, and their feature numbers are listed in Table

4.2.

Isotope Analysis of Bones/Teeth

All bones were sonicated for 5 minutes and then vacuum-dried for two days in preparation for grinding. The bones were then ground using a mortar and pestle and run through multiple sieves; the portion ranging in size from 250 to 1000 µm was collected for collagen extraction. This fraction was weighed and then distributed evenly on a layer of glass wool in a Pyrex coarse frit glass filter funnel. Hydrochloric acid (0.2M for well-preserved bone, 0.1M for poorly-preserved bone) was added to the funnel to demineralize bone. This hydrochloric acid was refreshed every twelve hours until the bone fragments stopped bubbling and became translucent. The funnels were then rinsed 6-8 times with deionized water, and then filled with 0.125M NaOH (0.625M for poorly-preserved samples) and allowed to soak for at least 20 hours. The funnels were rinsed again 6-8 times with water, then filled with 10^{-3} M HCl. The funnels were then loaded into a gravity oven at 70°C.

After 5 hours in the gravity oven, 100 µl of 1M HCl was added to the funnels and any evaporated 10^{-3} M HCl was replaced. Samples were heated in the gravity oven at least overnight, until the collagen was completely dissolved. The samples were then transferred to open Erlenmeyer flasks and continued to heat in the gravity oven until they had concentrated down to ~2 mL. Samples were then transferred to weighed scintillation vials and further concentrated to roughly 1 mL in volume. The samples were then removed from the gravity oven and placed in a freezer for 3 hours before being transferred to a vacuum freeze-dryer for 2-3 days. When the collagen completely dried, the vials were weighed and the yield calculated.

The Carlo-Erba NC2500 Elemental Analyzer at the Illinois State Geological Survey was used to analyze the $\delta^{13}C$ and $\delta^{15}N$ content of the samples. The dried collagen was homogenized and a minimum of 800 µg of collagen powder (a higher mass was used for samples with low collagen yield) was weighed into a tin foil capsule for analysis. Up to 2 mg of sample was analyzed if a low collagen yield was obtained.

A total of 34 samples in compressed tin capsules were placed in the Elemental Analyzer carousel, along with 14 samples of standards of thiourea, L-serine and hydroxyl-L-proline. The elemental analyzer converts organic matter into purified N_2 and CO_2 and transfers these gases in a helium carrier gas controlled by a Thermo-Finnegan ConFlo IV device to the Finnegan MAT 252 isotope ratio mass spectrometer. Carbon and nitrogen isotope ratios are reported as $\delta^{13}C$ and $\delta^{15}N$ relative to the PDB and AIR standards, respectively. Precision of analysis is generally ± 0.1‰ for $\delta^{13}C$ and 0.2‰ for $\delta^{15}N$.

Bone and enamel apatite were extracted using methods based on Balasse et al. (2002) . Roughly 15 mg of the bone fraction ranging in size from 37-63 μm was used. To remove bone proteins (mainly collagen), the bone was treated with 2.6% NaOCl (Clorox bleach) and allowed to sit overnight. The bleach was decanted, and then refreshed, and the sample sat for an additional day. The bleach was decanted again, and then the bone was rinsed four times with distilled water and 0.1 M acetic acid was added. The sample sat for four hours, and then was rinsed four times with distilled water, and was then freeze-dried.

All the teeth used in this study were lower M3s, to limit age affects that might alter the isotope values. Prior to removing enamel, the tooth was cleaned with a carbide drill tip to remove any surface soil. A scalpel was used to separate the enamel from the tooth, and a diamond-tipped drill bit was used to remove any dentine that remained attached to the enamel. The enamel was ground using an agate mortar and pestle, and 8 mg of crushed enamel was weighed out for extraction. The protocol for bone was followed for the enamel apatite, except that bleach was only added once, for one day, because enamel has much less organic matter.

The Kiel III automated carbonate reaction device at the Illinois State Geological Survey was used to convert apatite carbonate to CO_2 by reaction with 100% phosphoric acid under vacuum, and purify the gas by cryogenic distillation. Purified CO_2 is transferred to the Finnegan MAT 252 isotope ratio mass spectrometer to analyze its stable carbon and oxygen isotope ratios. The dried apatite was homogenized and 500 μg of bone or 700 μg of enamel was weighed out into glass vials. Each carousel included 38 apatite

samples, and three samples of National Institute of Standards and Technology carbonate standards (NBS-18 and three of NBS-19, weighing ~40-60 mg). Results are reported as $\delta^{13}C$ and $\delta^{18}O$ values relative to the V-PDB standard and V-SMOW standards (International Atomic Energy Agency, Vienna, Austria). Precision of analysis on the MAT 252 is generally ± 0.06‰ for $\delta^{13}C$ and 0.1‰ for $\delta^{18}O$.

A formula was used to calculate the percent C4 in the dogs' diet, using the following equations (Schwarcz, 1991; Ambrose et al., 1997):

$$\text{Collagen: \%C4} = -25 - (\delta 13C - 5.1‰)/15 * 100$$

$$\text{Apatite: \%C4} = -25 - (\delta 13C - 9.4‰)/15 * 100$$

The collagen $\delta^{13}C$ and $\delta^{15}N$ values of the dogs were compared to the values of humans from 15 Mississippian sites, both from the American Bottom and from more outlying areas, and with ages ranging from the Early to Late Mississippian periods (Bender et al., 1981; Katzenberg, 1989; Schober, 1998; Schoeninger and Schurr, 1998; Hedman et al., 2002; Balasse et al., 2003; Emerson et al., 2005). A full list of the sites used and their isotope means and standard deviations can be found in Table 4.5.

DNA Extraction of Coprolites

Prior to amplification, all extraction steps involving DNA were performed in the Ancient DNA lab at the Carl R. Woese Institute for Genomic Biology at the University of Illinois, which is physically separate from any lab wherein modern DNA samples are handled. In this lab, all workers wear full-body coverings and multiple layers of gloves to prevent sample contamination, and all surfaces and tools are cleaned with bleach and DNA-Off before and after use, and a UV Cross-linker is used to treat all lab equipment. DNA from ancient samples is degraded and fragile, and is easily contaminated with DNA from modern samples (Willerslev and Cooper, 2005). These precautions are taken to limit contamination as much as possible.

Coprolite samples were weighed prior to sampling. Using a scalpel, the end of the coprolite was sliced off, exposing the interior. A Dremel drill was used to drill the inside of the coprolite for 300-400 mg powder, which was extracted using the QIAmp DNA Mini Stool Kit following the protocol in Archie et al

(2003), with the following additional modifications: the powder was dissolved in 800 uL EDTA and 300 uL N-lauryl sarcosine, tubes were vortexed for 5 minutes with the InhibitEx tablet, and the DNA was eluted using two 30 uL aliquots of AE buffer, with 30 minutes of incubation time for each aliquot. A dog bone control from the same archaeological site was co-extracted with one of the extractions, to screen for lab and environmental contamination. A second control was used, which was a dog bone extracted using a different method, to differentiate between lab reagent contamination and environmental contamination. Extracts from dog bones should only contain dog DNA, so any other species identified would be considered a contaminant. If a contaminant was found in both controls, it was likely due to incomplete decontamination of the sample, or to contamination of lab reagents, and in either case the taxa was removed from analysis if it was identified in both controls and samples.

Bacteria in soil and in feces contain enzymes that can inhibit the polymerases in PCR, which is a common problem for ancient DNA samples (Alaeddini, 2012; Kemp et al., 2014). The extracted coprolite DNA was tested for PCR inhibition by preparing a PCR reaction using DNA extracted from ancient dog bones, and dog mitochondrial DNA primers (Witt et al 2015). The coprolite extracts were then used to "spike" the dog DNA samples and amplification was attempted. If the sample spiked with a coprolite extract failed to amplify (which would indicate PCR inhibition), the extract was run through a silica column using a Qiagen PCR Purification Kit. This step was repeated until the spiked control amplified successfully, which indicated that the majority of PCR inhibitors was removed. All coprolites that were submitted for sequencing took no more than 2 re-extractions to remove all inhibitors.

Libraries were built from the extracts using the NEBNext Ultra DNA Library Prep Kit for Illumina. The library prep was cleaned to remove adapters using the Agencourt Ampure XP beads, and samples were indexed using the NEBNext Multiplex Oligos for Illumina prior to amplification, which was done as recommended by the manufacturer for 12 cycles. A second amplification was prepared from the amplified product using Phusion High Fidelity Master Mix with a reduced volume of primers (1.5 uL instead of 2.5),

the addition of DMSO and 1 uL of BSA, with 5 uL of DNA, with four reactions prepared per sample. Thermocycling conditions followed the directions of the manufacturer, using the maximum time for the denaturing, annealing and extension steps and a 65°C annealing temperature, for 12 cycles. The four reactions per sample were pooled and cleaned using a Qiagen Minelute PCR Purification Kit. The samples were visualized on an agarose gel and quantitated using a Qubit Fluorimeter. The samples were pooled and shotgun-sequenced on two lanes of an Illumina HiSeq 4000, at the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign.

Once sequencing results were obtained, the sequencing reads were processed using a bioinformatics pipeline. Adapters were trimmed from the reads using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html), and a hard trim of 3 base pairs (bp) and a soft quality trim was also applied to the reads. Only reads that were greater than 50 bp in length were retained for analysis. The reads were de-duplicated, and then assembled into contigs de novo using Abyss version 1.3.4 with a k-mer length of 36 (Simpson et al., 2009). Using Abyss, each read is split into all possible 36-nucleotide kmers, and then they are assembled into larger contigs where there is overlap between the reads using a de Bruijn graph data structure. This is repeated for three iterations, to maximize contig assembly. The goal of this assembly was to combine the shorter reads into longer fragments, in hopes of getting more specific matches to modern taxa. The assembled reads were then compared to animal and plant DNA organellar DNA using a BLAST search. Two BLAST databases were used in the analysis, both of which were subsets of the RefSeq database, a non-redundant database of DNA sequences from the National Center for Biotechnology Information (NCBI). The mitochondrial DNA database included all taxa from Metazoa and Embryophyta, and the plastid database included all taxa from Embryophyta. Organellar DNA is more abundant in cells than autosomal DNA, so it is likely that there will be more mitochondrial and chloroplast DNA present than other sources of DNA (Alonso et al., 2004; Schwarz et al., 2009; Allentoft et al., 2012). The trimmed, de-duplicated and assembled reads were searched against each custom BLAST

database using blast+ 2.6.0 (Camacho et al., 2009), and the BLAST results were viewed in MEGAN6 Community Edition to determine the taxonomic identity (Huson et al., 2016). MEGAN is a program that uses the BLAST output to classify reads taxonomically. One can look at the reads in a single sample, or compare them between samples. MEGAN will classify taxa as specifically as it can, all the way down to species if possible. If the lower taxonomic levels are uncertain, MEGAN will classify it as a higher taxonomic level. Therefore, any ambiguous taxa that match with multiple families or genera can be excluded from the analysis.

The use of shotgun sequencing to identify dietary components in coprolites is fairly novel, and various guidelines have been suggested to distinguish between contaminant DNA and actual dietary DNA (Warinner et al., 2014; Pedersen et al., 2016). These guidelines range from a percent match requirement to the number of reads that must be present for a taxon before it can be accepted as present. For this study, three main requirements were chosen. First, any match identified using BLAST had to be a 90% or greater match to the read for it to be included in the analysis. Second, for a taxonomic family to be included in the analysis, three reads across all samples had to be identified as belonging to that family. The third requirement, regarding comparison to controls, had two levels. For the first level, “semi-strict”, if a taxon had an equal or greater number of reads present in the control when compared to the sample with the most reads, it was discarded. For the second level, “strict”, if a read representing the taxon was identified in the control, it was discarded. Both the Strict and Semi-Strict criteria were used, to determine what reads were present. Additionally, once a genus was identified, it was compared to lists of local flora and fauna in Illinois to determine if the species was likely to be real or was more likely to be the result of a database mismatch or contamination.

To further filter the results, we also took all reads that matched against the custom plant and animal databases and did a BLAST search against the nt database, which includes all nucleotide sequences uploaded to Genbank (which is part of NCBI). The drawback of using a specific database is that a sequence

may have a better match to a sequence outside of that database, and so the match within the database may be inaccurate. Additionally, if a read is from a region of the genome that is highly conserved, it is possible that the sequence may derive from bacteria. Because the database was limited to plant and animal DNA, bacterial DNA may appear to resemble the DNA of another organism. The nt database has DNA sequences across all kingdoms, so it is not biased in favor of any specific taxa. The reads that matched to mitochondrial DNA and the reads that matched to plastid DNA were separated for this analysis. The nt database was accessed on June 11th, 2017. The BLAST settings and sequence processing were the same as above, and these results were also visualized in MEGAN. The only difference in analysis is that we did not require at least three reads to be present, because there were so few reads left at the end of the analysis.

The primary goal of the dietary analysis was to identify maize. In addition to looking for maize in the BLAST results, we also used bwa to map the reads against the maize genome, and examined the ends of those reads using MapDamage (Jónsson et al., 2013). Ancient DNA typically shows damage on the ends, while modern DNA does not, so this is a sign to confirm the authenticity of the DNA. A large portion of the maize genome is conserved among related plants, and even if the reads that map to the genome do not specifically belong to maize, it is useful to determine if the plant DNA in general is ancient or just modern contamination.

Results

Isotope Analysis

All of the Janey B. Goode (JBG) samples analyzed yielded collagen or apatite, ranging from 1.52% to 15.03% collagen, 57.43% to 71% bone apatite, and 61% to 79% enamel apatite. The weight percent carbon, nitrogen, the $\delta^{13}C$ and $\delta^{15}N$ values, and the C:N ratio for bone collagen of each individual are all listed in Table 4.3, and the $\delta^{13}C$ and $\delta^{18}O$ values for the bone and enamel apatite are listed in Table 4.4. The weight percent carbon value was corrected slightly as the standards had only 87% of the expected value for weight percent carbon. The C:N ratio ranged from 3.05 to 3.25, with most values being within

.05 of the expected ratio of 3.21 (Ambrose, 1993). The $\delta^{13}C/^{12}C$ values for bone collagen ranged from -20.6‰ to -11.2‰ (mean = -16.02 ± 3.69‰), while the $\delta^{15}N/^{14}N$ values were much less variable, with minimum and maximum values of 8.4‰ and 9.7‰ (mean = 8.88 ± 0.46‰). The bone apatite $\delta^{13}C$ ranged from -4.7‰ to -12.0‰ (mean = -8.15 ± 2.59‰), and the enamel apatite $\delta^{13}C$ ranged from -4.0‰ to -8.2‰ (mean = -5.43 ± 1.96‰). The bone apatite $\delta^{18}O$ values showed low variance between individuals, ranging from 25.488 to 27.148 (mean=26.52 ± 0.48‰), while the enamel apatite $\delta^{18}O$ had greater variation, with a range from 22.813 to 29.59 (mean = 25.99 ± 2.64‰). The bone and enamel apatite $\delta^{18}O$ values show no correlation (p=0.629, Figure 4.5). A plot of the stable isotope values can be found in Figure 4.2, with the individuals color-coded by radiocarbon age. There is a weakly negative trend: an increase in $\delta^{13}C$ corresponds with a decrease in $\delta^{15}N$. In all but two individuals (JBG726-1 and JBG98-1), the enamel apatite and bone apatite $\delta^{13}C$ values are within 1.5‰ of one another. The collagen and bone apatite $\delta^{13}C$ show a strong correlation, with R^2=0.9917 (p=8.52 x10^{-7}, Figure 4.3), while the bone and enamel apatite $\delta^{13}C$ show no correlation (p=0.669, Figure 4.4). The bone collagen $\delta^{13}C$ is lower than the bone apatite $\delta^{13}C$ ($\Delta^{13}C$) by an average of 7.9‰. This interval has a high correlation with bone collagen $\delta^{13}C$ (p<.0001, Figure 4.7), but a low correlation with bone collagen $\delta^{15}N$ (p=0.472, Figure 4.6).

Higher $\delta^{13}C$ over time shows that the dogs increased maize consumption and decreased animal protein consumption from the Late Woodland to Mississippian periods. The oldest dog to show a significant percentage of maize in their diet (60% in collagen, 70% in whole diet) has a radiocarbon age of 991 ybp, suggesting that maize was well-established in the American Bottom by this time. The single individual that dates from the Terminal Late Woodland has a slightly higher $\delta^{13}C$ than the Late Woodland individuals, suggesting that maize may have been present in the American Bottom in small amounts around 1100 ybp, which is consistent with previous research (Ambrose, 1987; Vanderwarker et al., 2013; Simon, 2017).

Coprolite Sequencing

Maize (genus *Zea*) was identified in three of the eight coprolites. Coprolite 750-16 had 25 reads that matched to the maize mitochondrial sequence, coprolite 5146-2 had 3 reads that matched to the maize mitochondrial sequence, and coprolite 6227-15 had a single read that matched to the maize chloroplast sequence. Genus *Zea* was not found in either of the control sequences, and was also one of the few taxa to be identified from both chloroplast and mitochondrial DNA. The mapDamage plot for the maize reads (Figure 4.17) is jagged, as a result of the low sequence coverage, but does show a slight increase in damage on the ends of the reads, which is the damage pattern normally identified in ancient DNA. In addition to maize, a total of 82 genera (69 strict) were identified across 22 taxonomic orders using the custom RefSeq mitochondrial BLAST database (Figures 4.9 and 4.11), and a total of 73 genera (64 strict) were identified across 51 orders using the custom RefSeq chloroplast BLAST database (Figures 4.10 and 4.12). When these genera were compared to taxa that are found in Illinois, this narrowed the list of taxa to 6 animal and 4 plant genera (Table 4.7).

Nelumbo (American lotus), *Selaginella* (spikemoss), *Ulota* (a genera of moss) and *Orobanche* (broomrape) were identified in both controls and in nearly all of the samples, suggesting that they were contaminants. Many of these taxa were removed using the "semi-strict" or "strict" filtering methods. Other taxa were identified that were local to the American Bottom, and therefore were likely dietary components. The majority of these genera were fish, including *Couesius* (lake chub), *Lepomis* (a freshwater sunfish), and *Dorosoma* (gizzard shad). A number of other genera were also identified, including *Ardea* (great herons), *Cucurbita* (squash), *Solanum* (nightshade), and *Nicotiana* (tobacco). Some other local species were identified, but did not have 3 or more reads across all samples, including *Perca* (perch), *Helianthus* (sunflower) and *Atractosteus* (gar). Interestingly, two of the coprolites (2859-10 and 750-16) had reads from the genus *Toxocara*, a parasitic nematode that still infects dogs today. Even with the strict filtering criteria, however, a number of taxa remained that are non-native, including *Caesio*

(fusiliers, a tropical fish genus), suggesting that they are either contaminants to the extraction reagents or are incorrect matches from the BLAST search.

By using BLAST to compare only the reads that matched to my custom databases to the nt database, I found that the majority of the reads actually matched to bacteria (average: 86 %). A much smaller number of families and genera were identified (Figures 4.13-4.16), most of which were local species (Table 4.8). The identified taxa include Late Woodland crops such as squash and sunflower, as well as gizzard shad, hazelnut, wild grape, pawpaw, and amaranth. The only genera that could not be identified as local to Illinois were *Amborella* (a shrub from New Caledonia)*, Sesamum* (sesame)*, Cytinus* (a parasitic plant), and *Phoenix* (palm plants).

Discussion

Stable Isotopes

The variation in $\delta^{15}N$ in the sample was low, and the $\delta^{15}N$ value for all of the dogs is consistent with human pre-maize and post-maize populations in the American Bottom (Figure 4.8). These $\delta^{15}N$ values are higher than that of herbivores (4-7‰), but lower than that of carnivores (10-12‰), so the dogs likely ate low levels of animal protein (Fizet et al., 1995). Both Late Woodland and Mississippian dogs fall within the range of Late Woodland and Mississippian human $\delta^{15}N$, with the exception of the Ontario populations, which are both 2-3‰ higher than the other populations (Figure 4.8). The variation of $\delta^{13}C$ is much broader, and is consistent with a diet of both C_3 and C_4 plants, with higher numbers indicating more maize consumption and lower numbers indicating less (van der Merwe and Vogel, 1978). All of the Late Woodland dogs ate little to no maize, and a majority of the Mississippian dogs ate large amounts of maize. The Mississippian dogs (which all date to the Early Mississippian period) have a mean $\delta^{13}C$ value with populations from the Early and Middle Mississippian periods – in general the Late Mississippian populations had mean $\delta^{13}C$ values closer to -9‰. Late Woodland period dogs had 15% or less of their protein diet consisting of C_4 plants, while all Mississippian dogs dating to 991 ybp or more recent had over

50% of their protein diet from maize (Table 4.3). The dogs' $\delta^{13}C$ values are very similar to that of contemporaneous human populations, demonstrating that human and dog diets were likely very similar in the American Bottom (Figure 4.8). Their whole diet had even higher percentages of C_4 plants, with Late Woodland dogs consuming 40-50% C_4 plants and Mississippian dogs consuming as much as 80% C_4 plants (Table 3.4). There is an increase in maize consumption from the Late Woodland to Mississippian periods. The $\delta^{18}O$ values of the dogs are consistent with other carnivores and human populations (Bocherens et al., 1996; White et al., 2004). The Late Woodland individuals, on average, have a lower value than the Mississippian individuals, which suggests that the Mississippian individuals are getting more of their water from plants.

In the dogs from JBG, enamel apatite $\delta^{13}C$ is within 1.5‰ of bone apatite $\delta^{13}C$, with two exceptions (Table 4.4). In both cases, the bone apatite $\delta^{13}C$ is significantly lower than the enamel apatite $\delta^{13}C$. This suggests that those two dogs had different diets as puppies and adults, since enamel apatite reflects the diet at the time of tooth formation and bone apatite is reflective of diet as an adult. As puppies, the dogs likely had a diet with a higher percentage of C_4 plants, and then consumed less maize as adults. It is difficult to say whether those individuals were born at JBG and had a shift in diet as they grew, or if the dogs were born elsewhere, and then were transported to JBG, where they ate less maize.

The bone apatite shows an average offset from bone collagen of 7.87‰ ± 0.42. This is higher than the amount of difference typically found in carnivores and other animals with monogastric, non-methanogenic fermentation digestive physiologies (Ambrose and Norr, 1993). This indicates that the whole diet and the protein component of the diet have different C_3/C_4 ratios. Given the low $\delta^{15}N$ values, the dogs were likely eating a mainly plant-based diet, and the collagen offset from diet should be roughly 4.5‰ (Ambrose et al., 1997). The higher offset for this population indicates that the protein component of the dogs' diet had a lower $\delta^{13}C$, and likely derived from both C_3 and C_4 sources, while the maize in their diet would have increased the $\delta^{13}C$ value for the whole diet. The coprolites that contained maize DNA also

had DNA from native fish, suggesting that consumption of fish is responsible for the decrease in $\delta^{13}C$ in the protein aspect of the diet.

In addition to the dietary isotope values, the radiocarbon ages of these individuals may also be affected by the fact that the dogs (and presumably the humans) from JBG were eating fish. Lakes and rivers can contain large amounts of dissolved inorganic (carbonate) carbon from ancient limestone rocks and organic carbon with older radiocarbon ages (the Reservoir Effect), which can affect radiocarbon dating and make samples look older than they actually are, in some cases as much as 2000 years older (Geyh et al., 1999). This is known as the reservoir effect, and can be calculated by determining the radiocarbon age of modern fish or aquatic plants, but it is not a perfect measure, because the estimate of the radiocarbon reservoir changes over time, and is affected by precipitation (Philippsen, 2013). By calculating the Reservoir Effect in this area of the Mississippi River, it would help to cement the timing of maize consumption in the American Bottom.

Dietary Components

Based on the carbon isotope results (Figure 4.2) and the coprolite sequencing results (Table 4.7), we successfully identified maize in the diet of the dogs from Janey B. Goode. The presence of maize in three coprolites confirms that the high $\delta^{13}C$ in the Mississippian dogs is due to maize, not another C_4 plant. This is not a surprising find, given that few native plants in the region are C_4 plants, and that other C_4 plants that have been identified at American Bottom sites, such as panic grass (*Panicum spp.*) and amaranth (*Amaranthus spp.*) have only been identified in small numbers, suggesting that they contributed little to the overall diet of Mississippians (Bender et al., 1981; Vanderwarker et al., 2013). The map damage plots show some irregularity in damage pattern, primarily because of the low coverage of the maize genome in the coprolites – an example of the plot of DNA damage is shown in Figure 4.17. However, there is a slight increase in damage on the ends, showing that this DNA is ancient, and not contamination. Carbon isotope analysis of Janey B. Goode dog bones and teeth show that maize consumption was common in the

Mississippian period, while dogs from the Late Woodland period solely ate C_3 plants. Other plant species were also identified from the coprolites, including nightshade, tobacco, and squash. Additionally, multiple species of fish were identified, including freshwater sunfish and gizzard shad (Table 4.6). This is consistent with what is known about diet in the Mississippian and Late Woodland periods (van der Merwe and Vogel, 1978; Lopinot, 1997; Hedman et al., 2002; Smith, 2011), which included maize and a variety of other crops. The fish identified are all found in the Mississippi River, which runs very close to Cahokia. Macro-analysis of the coprolites revealed the presence of fish bones and scales, which also supports the DNA sequencing results (Fortier, 2015).

The identification of *Toxocara canis* in two of the eight coprolites suggests that some of the dogs at JBG were infected with this parasite. *T. canis* is a parasitic nematode that is transmitted through the soil – eggs are deposited in the feces of the dogs that are infected, and if dogs or humans consume the eggs, they can also be infected, which causes a disease known as toxocariasis (Despommier, 2003). This disease is most harmful to children, and can damage organs including the liver and lungs, and cause blindness. Toxocariasis is still common today, especially in children living in urban areas or from lower-income areas, and many puppies are infected with *T. canis* from birth (Robertson and Thompson, 2002). If the dogs at JBG were infected with *T. canis*, it is likely that some of the people living with the dogs had toxocariasis as well. *T. canis* was also identified in dogs associated with the Chiribaya culture in Peru (1300-600 ybp), and so there is a precedent for these parasites being present in ancient American dogs (Richardson et al., 2012).

Some taxa that have been identified in other studies as local to the American Bottom are not found in these coprolites. Deer and other woodland animals were thought to be a part of the Mississippian diet, along with waterfowl and the crops that form the Southeastern Agricultural Complex (Kelly, 1997; Lopinot, 1997; Hedman et al., 2002; Yerkes, 2011). The remains of numerous plant and animal taxa have been recovered from Late Woodland and Mississippian contexts (Table 4.6), but only a fraction of them

have been identified from these coprolites. Some of these absences may be due to limitations of the database. For example, at NCBI, which is the primary resource for DNA sequence data, there is only a single mitochondrial DNA sequence for *Chenopodium*, and it is only 230 base pairs (bp) long. In the RefSeq database, which was used for the BLAST searches, only one species from the family of *Chenopodium* (*Amaranthaceae)* was included, but there was a single read which matched it, suggesting that chenopods may have been present in the coprolites. Similarly, the only frog genus identified in the coprolites was *Anomaloglossus,* which is a poison dart frog native to Brazil. However, *Anomaloglossus* is part of the same superfamily as many species of toad native to Illinois, none of which are found in the RefSeq database that was used. In other cases, the species one might expect to find in the coprolites may be present, but in such small numbers of reads that they were removed from further consideration. *Helianthus* (sunflower) and *Atractosteus* (gar) were present in one coprolite, but in only one read each, so they failed to pass the filtering criteria.

Some of the limits of our findings may be due to the coprolites themselves. Heat denatures DNA, and so cooking would likely destroy much of the DNA in meat the dogs were eating. It is therefore likely that there is a bias towards raw foods the dogs were eating, rather than cooked. It must also be noted that coprolites likely only reflect the most recent meals, so if something was out of season when the coprolite was produced, it would not be found in the diet. Only a small number of coprolites that were likely produced at JBG were preserved, and certain seasons might better facilitate the preservation of dog feces, leading to a bias in their contents based on seasonality (Fortier, 2015). Given the small number of species that have been identified from the coprolites, however, it is difficult to assess whether the coprolites come from a particular season.

Pitfalls of Shotgun Sequencing

While the analysis of coprolites has yielded interesting results in regard to dog diet in the American Bottom, there is more analysis to be done. Shotgun sequencing amplifies everything in a sample,

and it is a useful tool when trying to identify multiple species that may only be present in small amounts. However, shotgun sequencing produces millions of reads per sample, and a majority of those reads belong to bacteria and other contaminants. Distinguishing between contamination and "authentic" reads can be a challenge.

In addition to the absence of some taxa that might be expected in the coprolites, there were many taxa identified that were unexpected. These included large numbers of sponge and moss reads. The genera that showed up in the largest numbers include *Ulota* and *Nyholmiella,* (both mosses), *Nelumbo* (a lotus), and *Orobanche* and *Phelipanche* (both broomrape). Given that these samples were found in both controls, they are likely to either be environmental contaminants in the soil that the bones and coprolites were recovered from, or contaminants of the library reagents used. Contamination of lab reagents is problematic for ancient DNA research, especially when shotgun sequencing is used to identify trace numbers of reads associated with microbiome or dietary research. Even if a small number of contaminant reads are present in a reagent, they still have the potential to be amplified and then sequenced. DNA from a number of animals, including cows and pigs, is known to be present as contaminants (Leonard et al., 2007), along with a number of microbes (Salter et al., 2014), but less is known about the plants that contaminate lab reagents. Another source of contamination, which is harder to identify, is the environmental DNA found in the soil with the coprolite. Attempts were made to sample the inside of the coprolites only, to limit this concern, but the coprolites could not easily be decontaminated without destroying them, and so it's possible that environmental DNA was sampled in addition to the dietary DNA that was identified.

By searching all mitochondrial and chloroplast matches against the NCBI nt database, most of the reads belonging to non-local taxa were eliminated, leaving only four genera that are not local to Illinois (Table 4.8). Most of the reads identified in the initial BLAST search matched to bacterial taxa, suggesting that those reads derive from highly conserved regions of the genome. Some of the local taxa, including

lake chub, freshwater sunfish, and maize, were not found in the second BLAST search. However, the families of some of the taxa (Poaceae for *Zea*) were still identified, suggesting that the reads matched to multiple taxa within Poaceae. These species may not all be local to Illinois, although further investigation is needed to determine what the possible identified matches were.

Another concern with correctly identifying the taxa present in a coprolite is in regard to DNA damage. In ancient DNA samples, the DNA fragments can be damaged by exposure to water, oxygen, sunlight, heat, and bacterial digestive processes (Gilbert et al., 2003; Dabney et al., 2013). This damage can cause strand breakage (leading to further fragmentation), base pair removal, and even changes to the DNA sequence itself. The most common of these "miscoding lesions" is cytosine deamination, in which a cytosine is changed to a uracil, which is then amplified as if it were a thymine (Gilbert et al., 2003). The reads that were used in this analysis were only 50-97 base pairs long, and if damage caused changes to the DNA sequence, it's possible that this would cause a mismatch between the read and the organism it matched most closely. Even one damaged nucleotide would be 1-2% of the sequence used, and the read might match to a taxon in another genus, or even another family. Some of these taxonomic assignments from BLAST may be incorrect due to accumulation of DNA damage.

Further investigation of these samples, especially the coprolites, is needed, given that this research is still very exploratory. The coprolites should be radiocarbon dated, to provide additional information about when these dietary components were actually being consumed. By dating the coprolites containing maize, it will be possible to determine when maize became a large part of the diet of the dogs at Janey B. Goode, and this can be compared to the bone isotope results, to see if the increase in maize consumption was very gradual or very abrupt. If the coprolites date to 1010 AD or younger, it would suggest that maize had a very rapid increase in consumption. If instead some of the coprolites are older, it indicates that maize was being consumed in small quantities before the Mississippian period, and the amount of maize in the diet slowly increased.

Conclusions

The dogs at Janey B. Goode provide insight into the Late Woodland-Mississippian transition in the American Bottom, especially in regard to changes in diet during this time. Isotopic analysis of dog bones and teeth indicate that during the Late Woodland the dogs ate almost no maize, while during the Mississippian period they had a diet with large amounts of maize. DNA sequencing of dog coprolites indicates that dogs had a varied diet of maize, squash, tobacco, heron, and fish. An internal parasite, *T. canis*, was also identified, which also likely affected humans at Janey B. Goode. The sequencing results also show a few taxa that are non-native, suggesting that further filtering of the genera identified is needed to get a better view of dog diet during the Late Woodland and Mississippian period. DNA sequencing of coprolites is a useful complement to dietary isotope analysis, as broad-scale trends can be narrowed down to specific taxa that were being consumed.

Figures and Tables

Figures

Figure 4.1: A map showing the location of the American Bottom in blue. Cahokia is denoted by the black square.

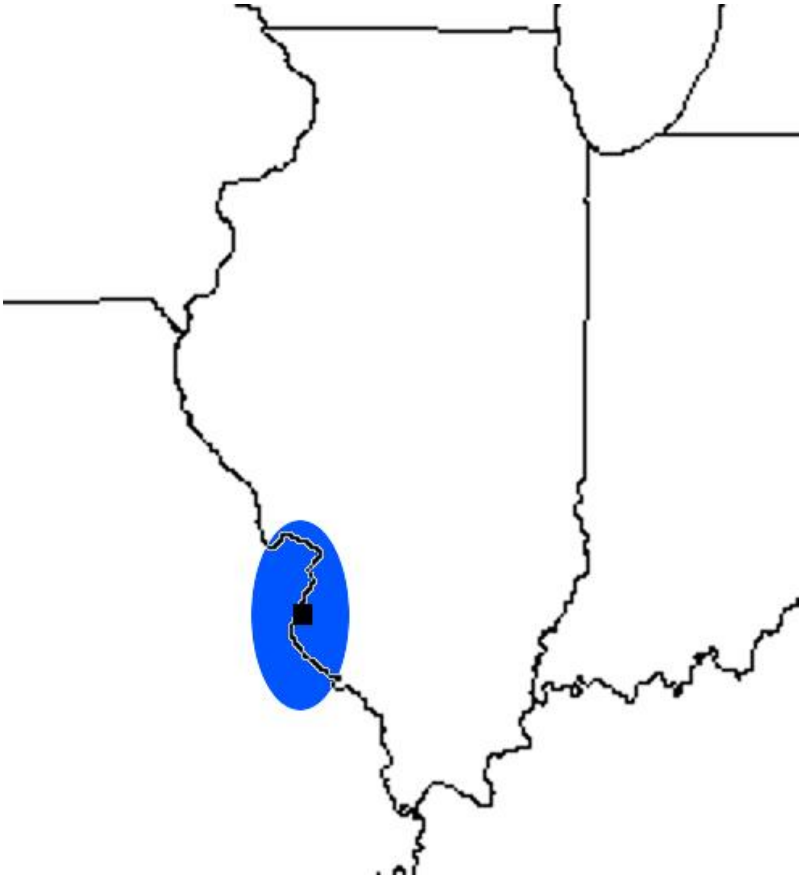


Figure 4.2: A plot of bone collagen isotope values for the dogs from Janey B. Goode. The samples are color-coded by time period – gray is unknown, orange is Late Woodland (>1100 ybp), pink is Terminal Late Woodland (1100-1000 ybp), and purple is Mississippian (<1000 ybp)

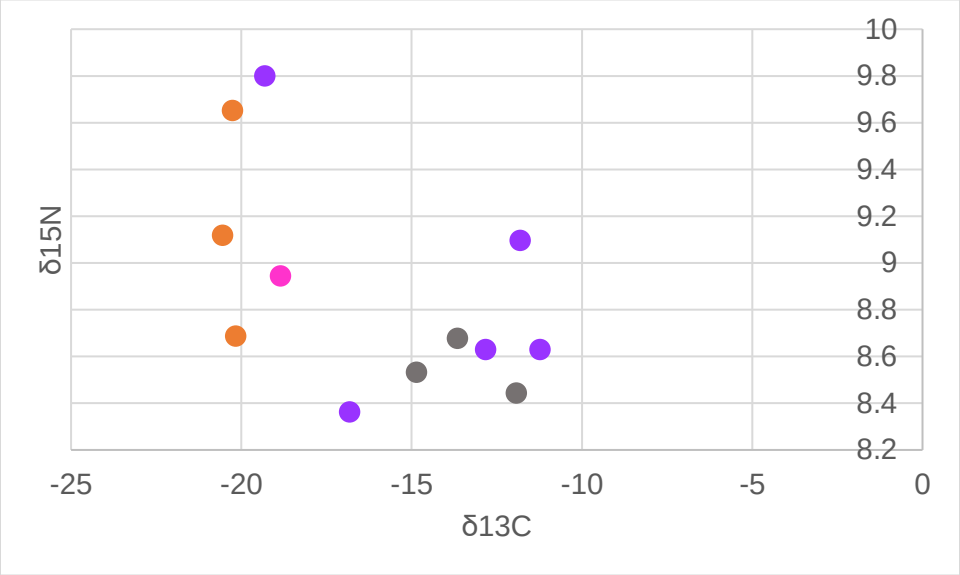


Figure 4.3: A plot comparing the correlation of bone collagen $\delta^{13}C$ to bone apatite $\delta^{13}C$ ($p=8.515 \times 10^{-7}$).

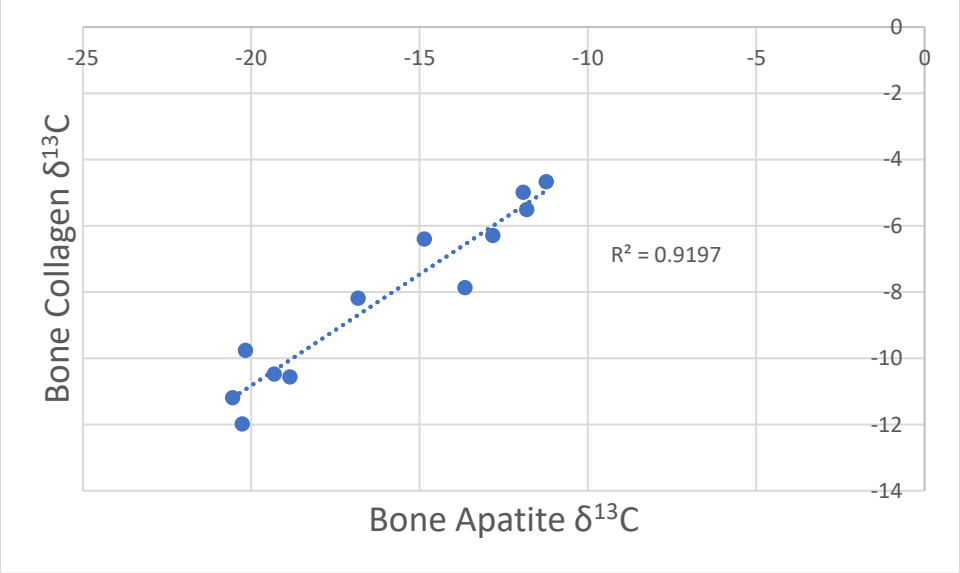


Figure 4.4: A plot comparing the bone and enamel apatite $\delta^{13}C$ values for each individual (p = 0.669).

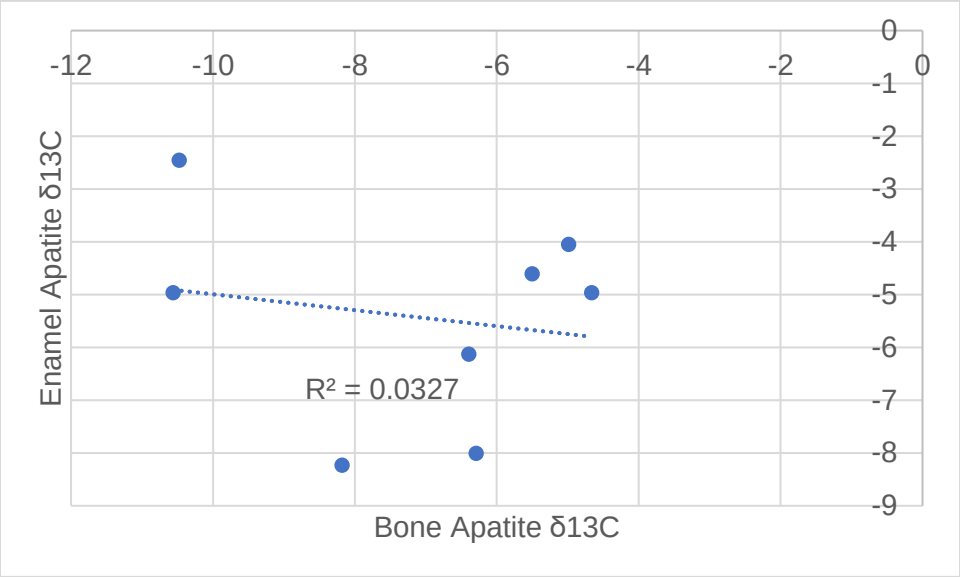


Figure 4.5: A plot comparing the bone and enamel apatite $\delta^{18}O$ values for each individual (p=0.629)

.

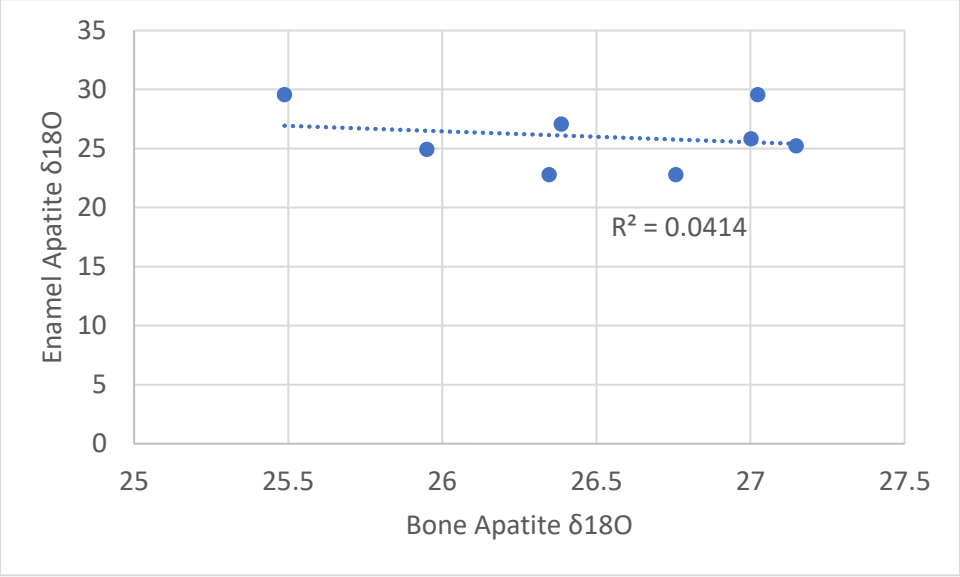


Figure 4.6: A plot comparing the $\delta^{15}N$ value to the difference in bone apatite and bone collagen $\delta^{13}C$ values for each individual (p=0.472).

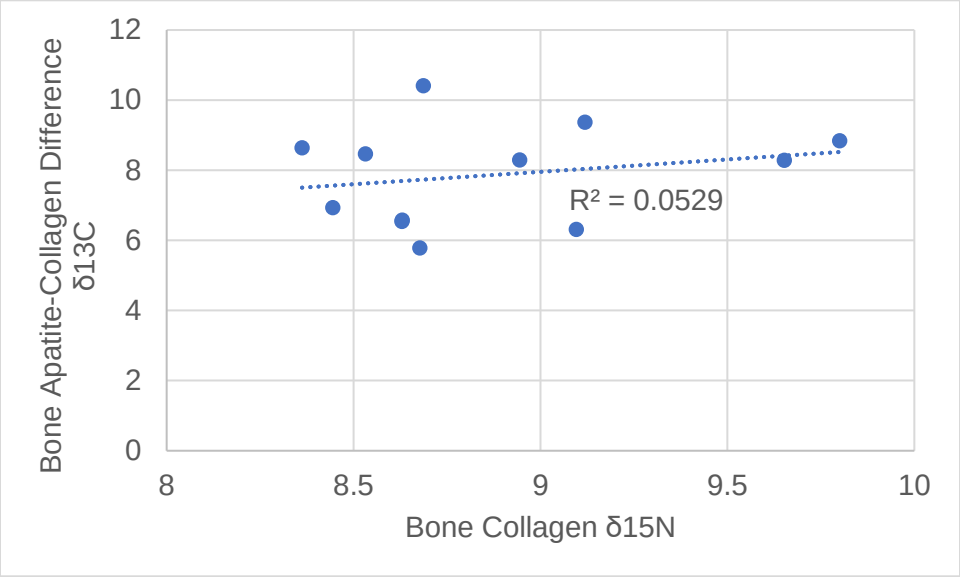


Figure 4.7: A plot comparing the bone collagen $\delta^{13}C$ value to the difference in bone apatite and bone collagen $\delta^{13}C$ values for each individual (p=0.000385).

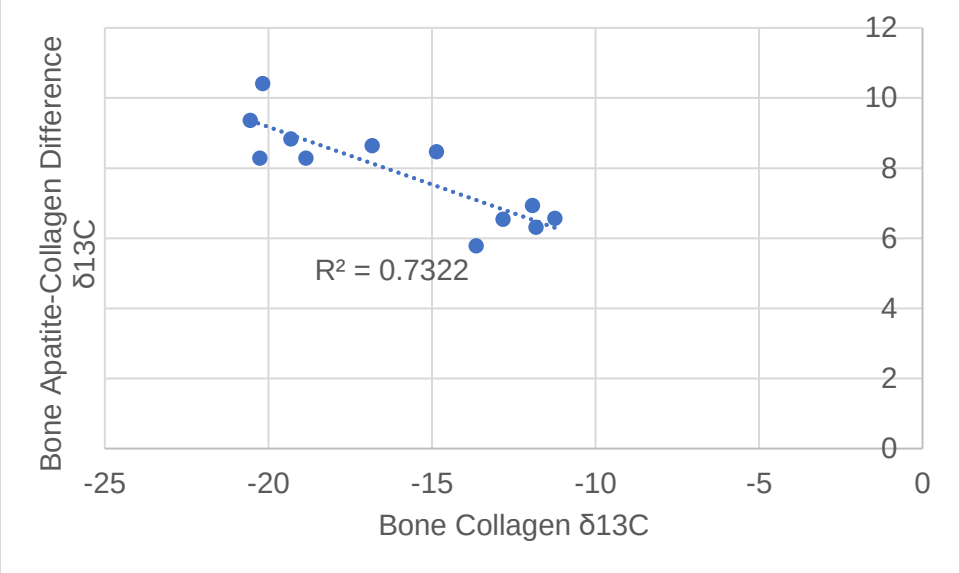


Figure 4.8: A plot comparing the collagen $\delta^{13}C$ and $\delta^{15}N$ values of the dogs analyzed in this study to human Mississippian populations, as listed in Table 4.7. Each dot represents a mean value for a different population, and the error bars show the standard deviation. Dogs are represented by squares, while humans are represented by circles.

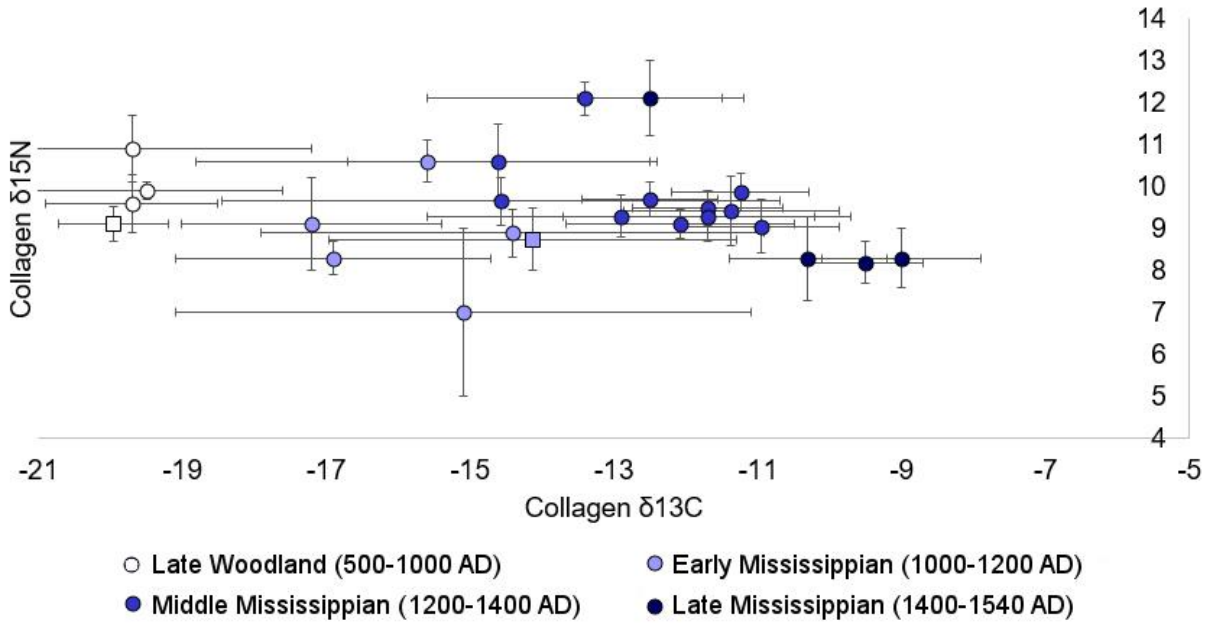


Figure 4.9: A plot of read counts for all of the genera found across all samples, using the mitochondrial BLAST database and the "semi-strict" filtering criteria. Any taxa that had the greatest number of reads in the control was removed from the analysis. The X axis is the genus names, in descending number of total reads identified, and the Y axis is the number of reads identified. Each colored bar is a different coprolite, while the gray bars are the two control samples.

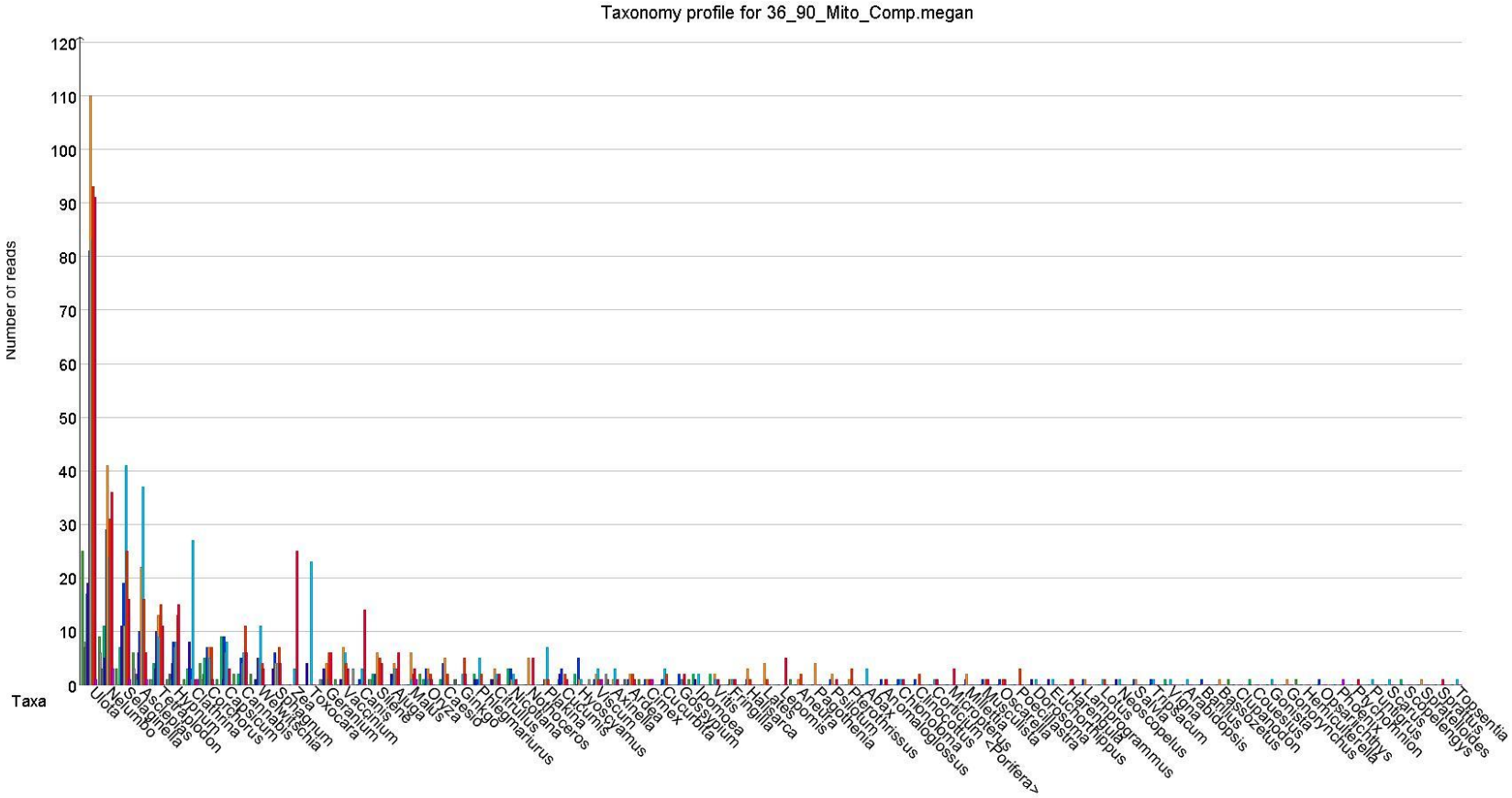


Figure 4.10: A plot of read counts for all of the genera found across all samples, using the chloroplast BLAST database. Any taxa that had the greatest number of reads in the control was removed from the analysis. The X axis is the genus names, in descending number of total reads identified, and the Y axis is the number of reads identified. Each colored bar is a different coprolite, while the gray bars are the two control samples.

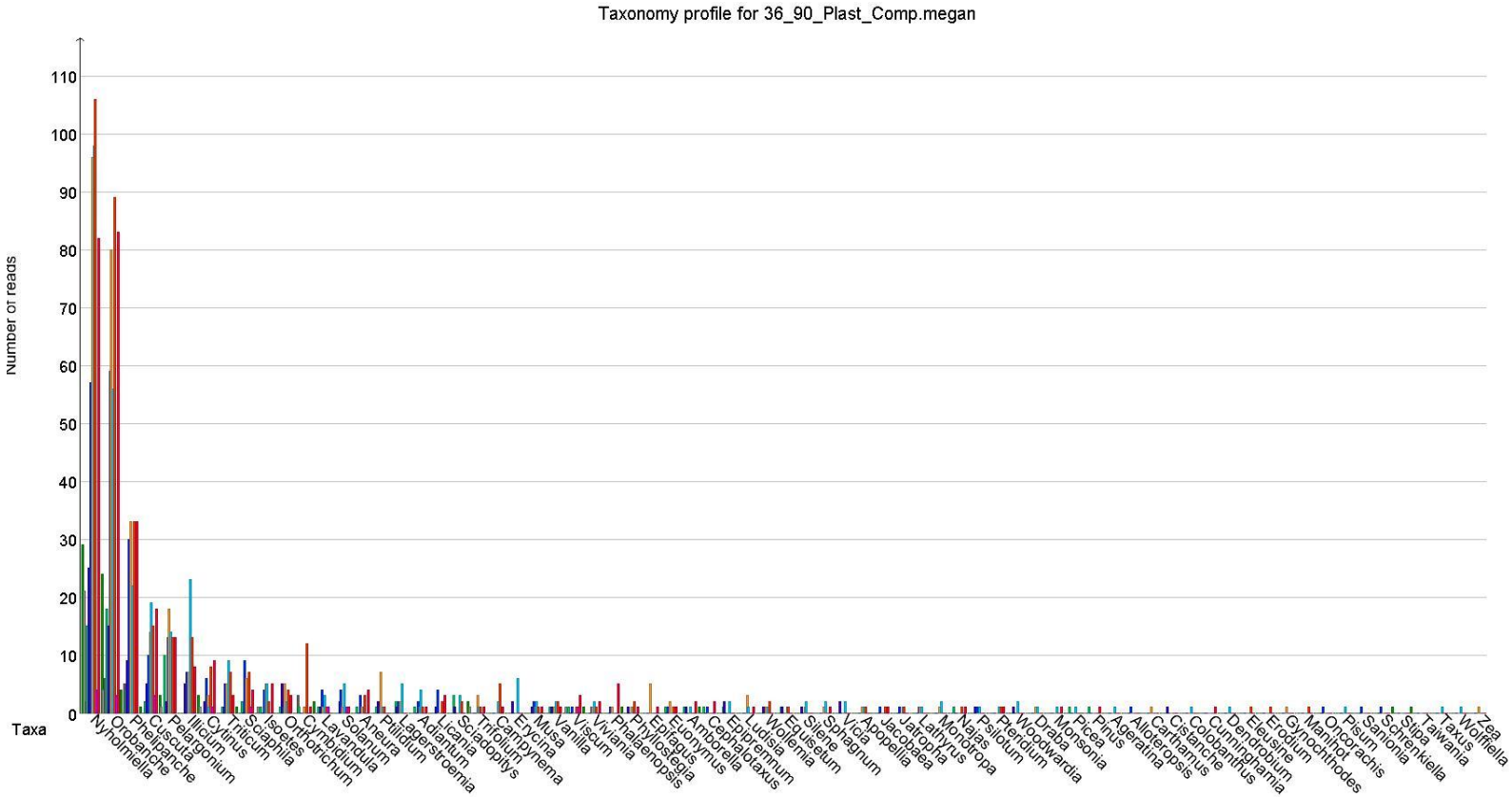


Figure 4.11: A plot of read counts for all of the genera found across all samples, using the mitochondrial BLAST database and the "strict" filtering criteria. Any taxa that had any reads identified in the control was removed from the analysis. The X axis is the genus names, in descending number of total reads identified, and the Y axis is the number of reads identified. Each colored bar is a different coprolite.

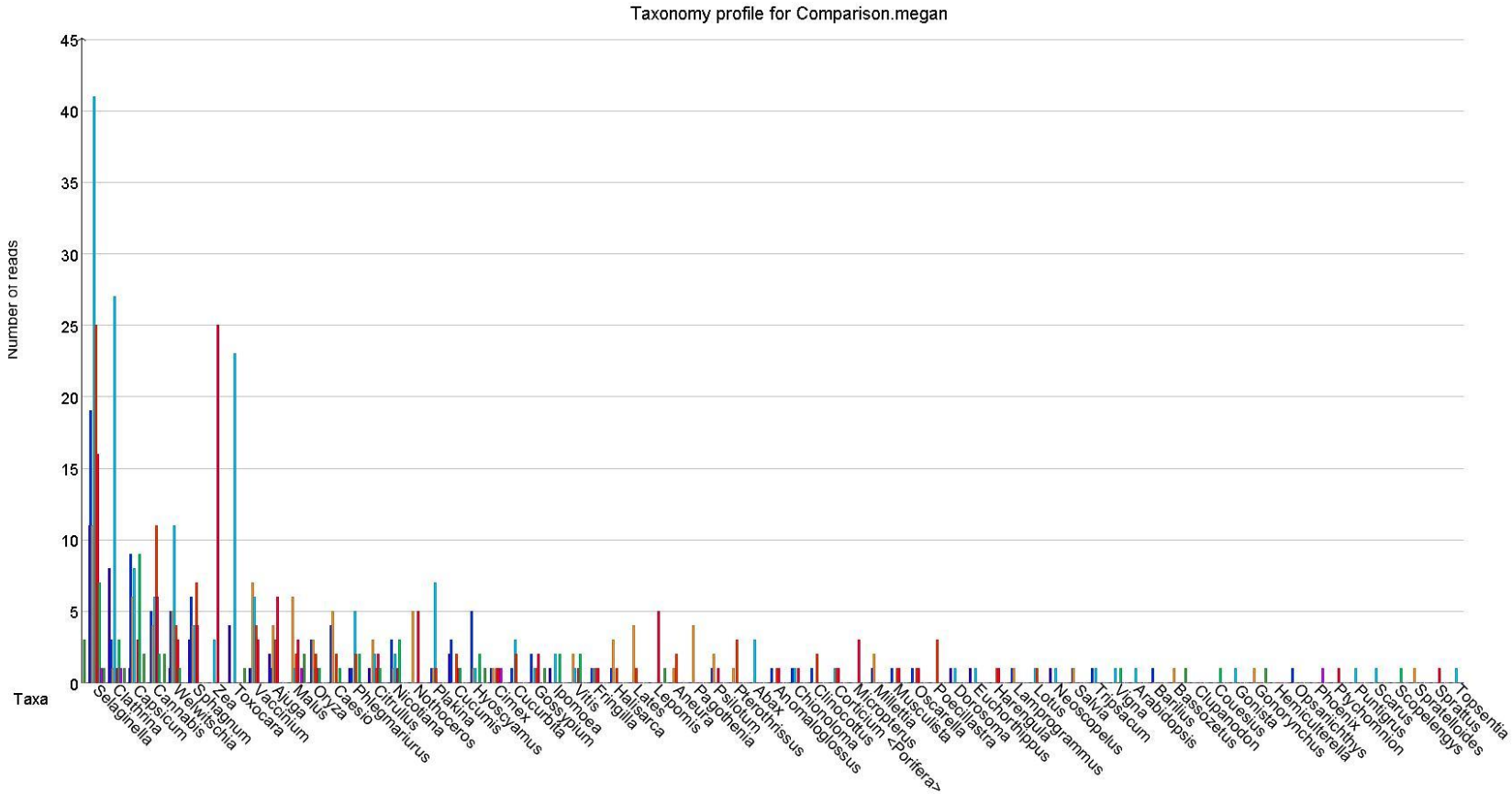


Figure 4.12: A plot of read counts for all of the genera found across all samples, using the chloroplast BLAST database and the "strict" filtering criteria. Any taxa that had any reads identified in the control was removed from the analysis. The X axis is the genus names, in descending number of total reads identified, and the Y axis is the number of reads identified. Each colored bar is a different coprolite.

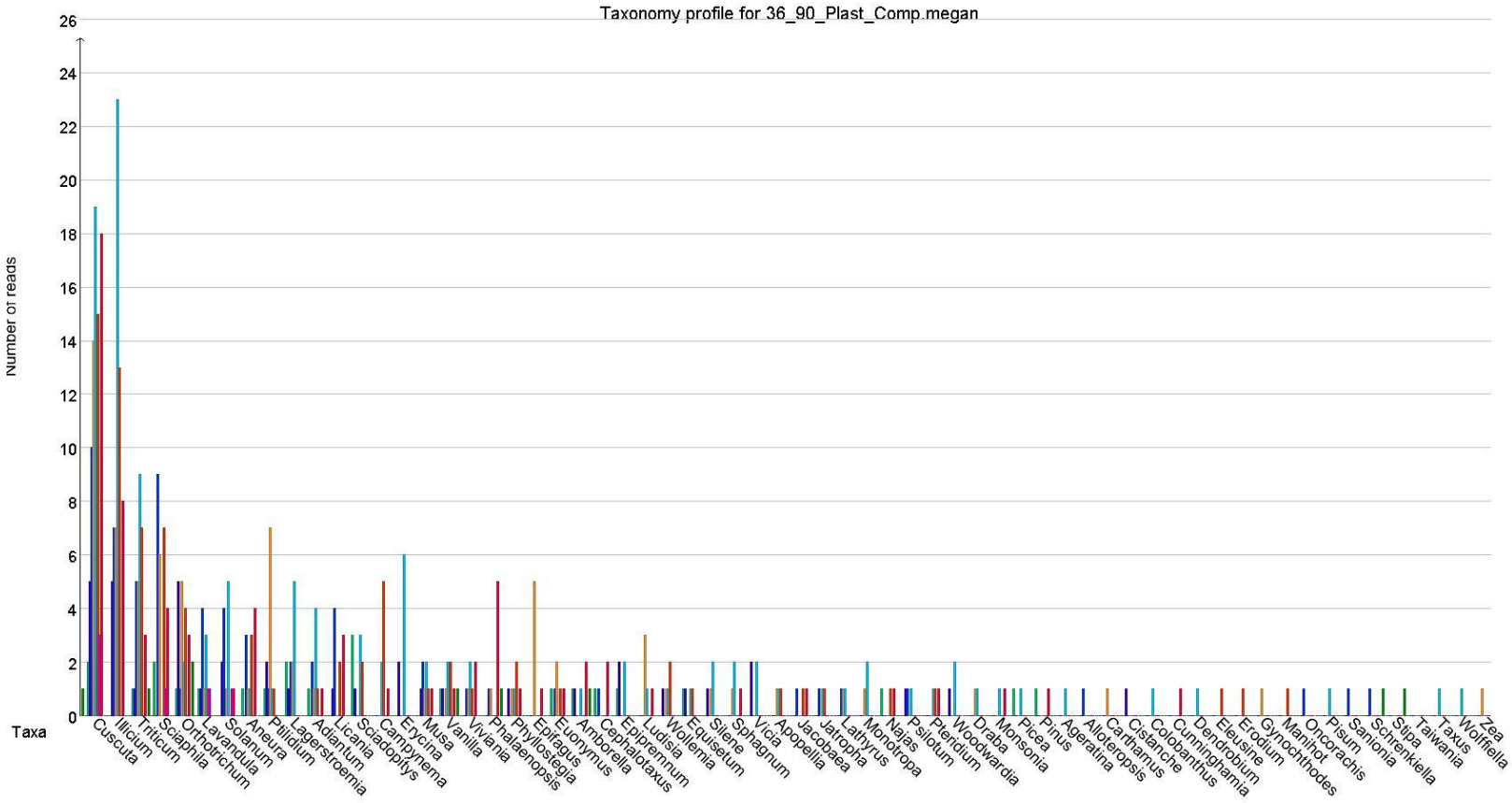


Figure 4.13: A plot of read counts for all of the families found across all samples, using the subset of reads that matched to the mitochondrial database searched against the nt database on NCBI. The X axis is the family names, and the Y axis is the number of reads identified. Each colored bar is a different coprolite, while the gray bars represent the controls.

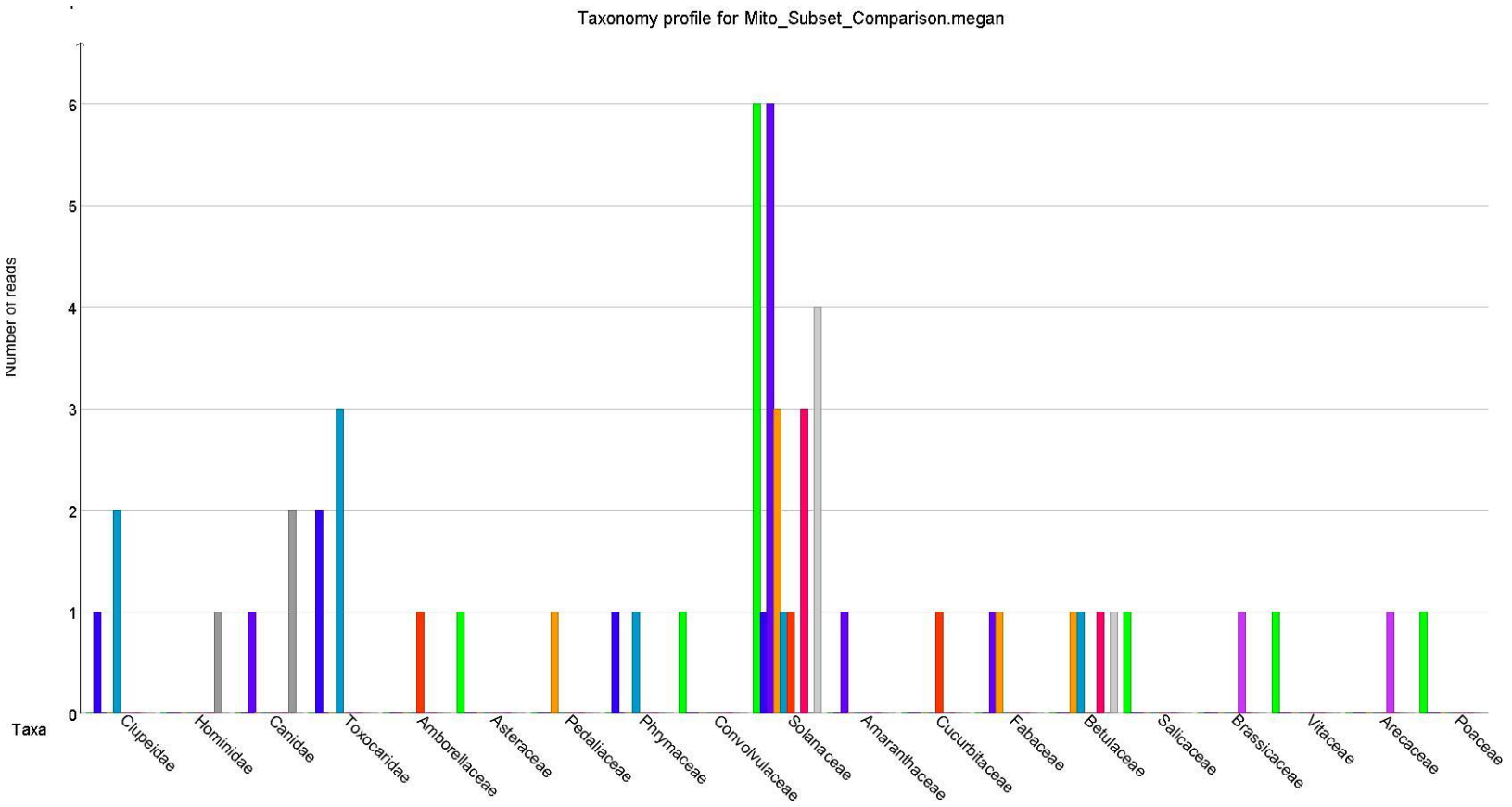


Figure 4.14: A plot of read counts for all of the genera found across all samples, using the subset of reads that matched to the mitochondrial database searched against the nt database on NCBI. The X axis is the genus names, and the Y axis is the number of reads identified. Each colored bar is a different coprolite, while the gray bars represent the controls.

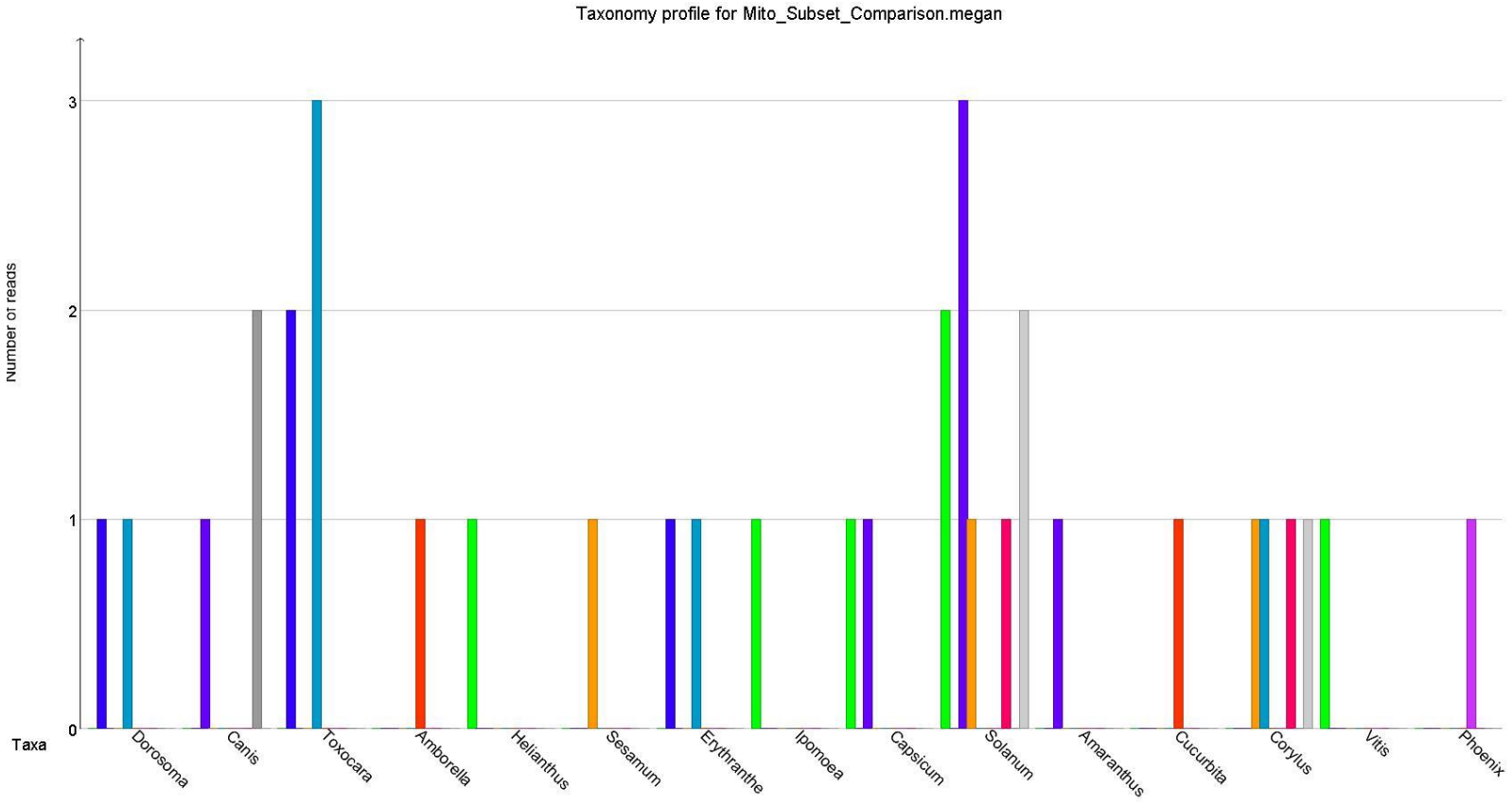


Figure 4.15: A plot of read counts for all of the families found across all samples, using the subset of reads that matched to the chloroplast database searched against the nt database on NCBI. The X axis is the family names, and the Y axis is the number of reads identified. Each colored bar is a different coprolite, while the gray bars represent the controls.

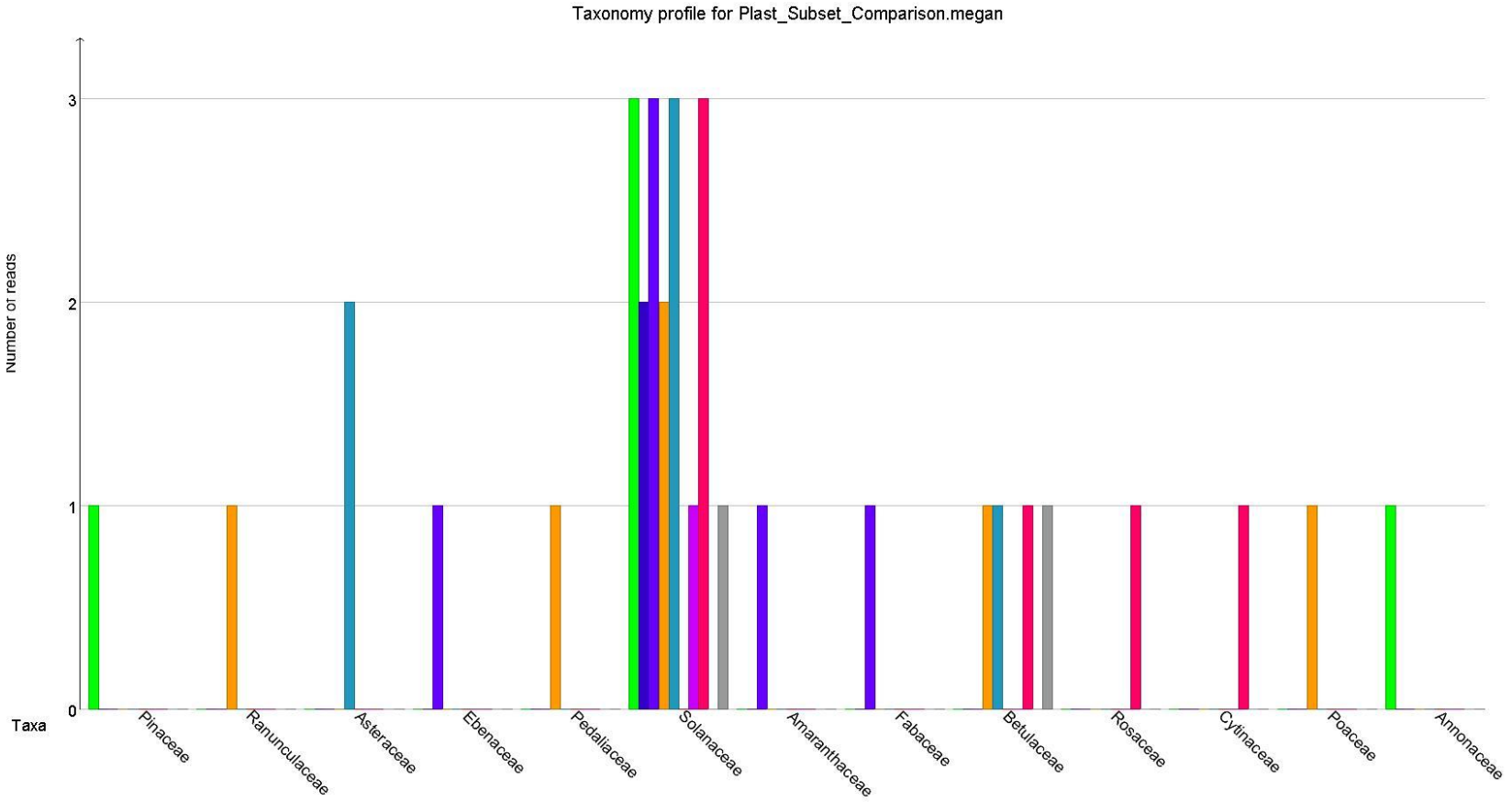


Figure 4.16 A plot of read counts for all of the genera found across all samples, using the subset of reads that matched to the mitochondrial database searched against the nt database on NCBI. The X axis is the genus names, and the Y axis is the number of reads identified. Each colored bar is a different coprolite, while the gray bars represent the controls.

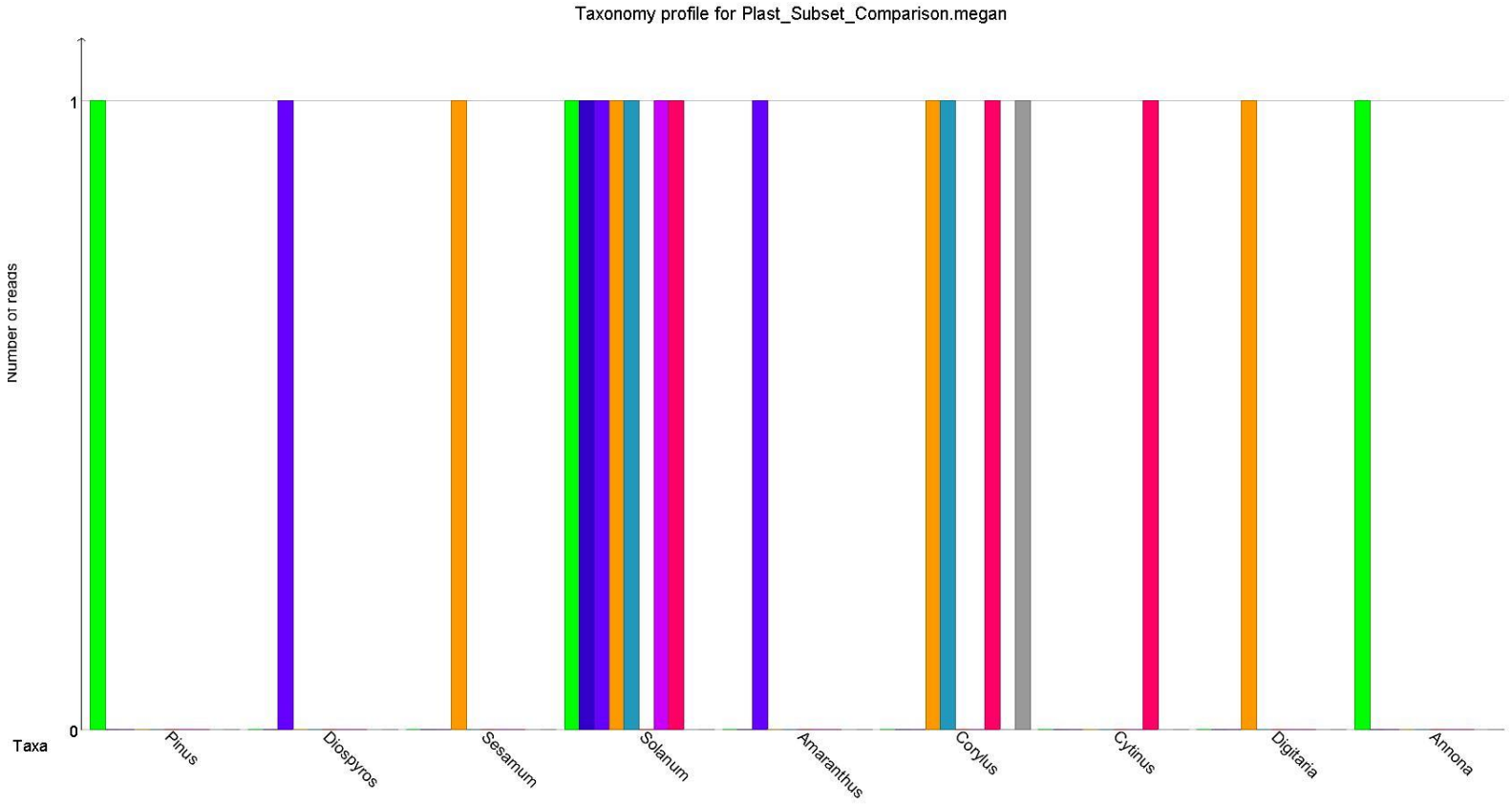
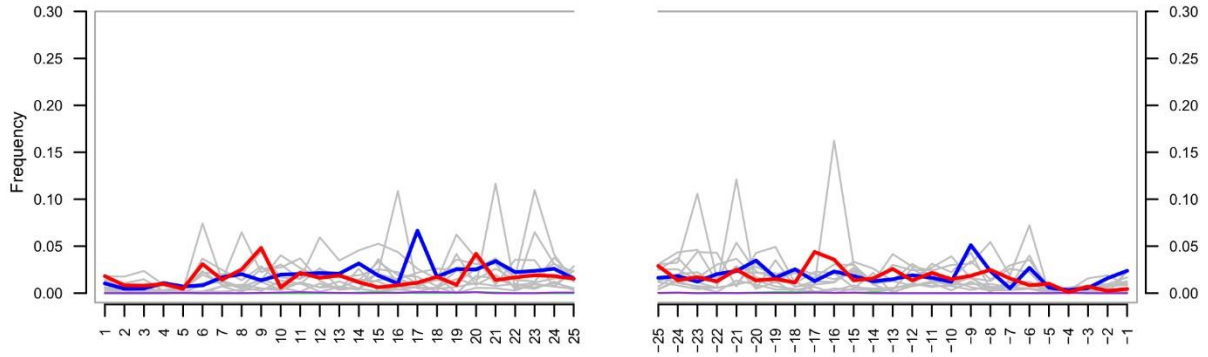


Figure 4.17: A MapDamage plot for the reads from coprolite JBG 750-16 which mapped to the maize genome. The red and blue lines track the rate of C->T transitions, and a higher peak shows a higher rate of damage. Here, the irregular damage pattern is due to a low rate of alignment.



Tables

Table 4.1: A list of the bones and teeth used in the isotope analysis. The feature number is listed, along with the estimated age using provenience, the radiocarbon age if known, whether a bone or tooth was sampled, and also the bone that was used. Radiocarbon dates have been provided by the Illinois State Archaeological Survey.

Individual	Suspected Age	Age (ybp)	Bone Sampled	Tooth Sampled	Bone Used
98-2	Terminal Late Woodland	987	X	X	Rib
274	Mississippian	1299	X		Humerus
635	Terminal Late Woodland	990	X	X	Rib
726	Late Woodland	1049	X	X	Rib
803	Late Woodland	Unknown	X		Rib
975	Terminal Late Woodland	Unknown	X	X	Rib
1030	Mississippian	1432	X		Rib
1337	Late Woodland	1484	X		Rib
1724	Terminal Late Woodland	991	X	X	Rib
3134	Late Woodland	999	X	X	Rib
5499	Terminal Late Woodland	914	X	X	Rib
6134	Late Woodland	Unknown	X	X	Radius

Table 4.2: A list of the coprolites used in the DNA analysis. The feature number is listed, along with whether or not it was sequenced, the number of raw reads, the number of reads after the trimming and de-duplication steps in the pipeline, and the number of de-duplicated and assembled reads used in the BLAST analysis, as well as the number of reads that matched mitochondrial or chloroplast genomes in the RefSeq databases. The * indicates a sample that was extracted using a different method.

Sample	Analyzed for DNA?	Raw Reads	Trimmed Reads	Assembled Reads	Hits to mtDNA	Hits to cpDNA
750-16	Yes	79,691,878	29,469,223	3,164,239	1243	1051
1045-13	Yes	83,314,584	26,206,075	3,110,977	916	817
2240-4	Yes	62,474,161	10,557,502	1,525,823	368	310
6227-15	Yes	72,375,269	40,565,183	3,325,616	1237	985
6553-2	No	-	-	-	-	-
1131-29	Yes	66,847,891	33,569,605	1,923,729	1083	965
2859-10	Yes	64,729,787	31,303,216	1,944,621	308	298
7186-2	No	-	-	-	-	-
7298-2	Yes	44,305,225	1,594,550	130,340	42	41
5146-2	Yes	102,139,995	42,358,981	2,752,170	1143	976
ESL166	Yes	40,609,814	7,705,867	583,487	285	241
JBG2 (control)	Yes*	8,131,302	7,525,498	196,496	208	120
JBG47 (control)	Yes	57,091,326	2,476,144	251,131	107	75

Table 4.3: A summary of the stable isotope results for the collagen samples. All radiocarbon dates provided by the Illinois State Archaeological Survey

Individual	Radiocarbon Age (AD)	ID	Wt% C	$\delta^{13}C$	Wt% N	$\delta^{15}N$	% C4 diet (protein)
JBG6134-1	N/A	KEW1	3.40	-11.924	1.49	8.444	59.2
JBG5499	1086	KEW2	32.70	-11.235	13.82	8.63	63.5
JBG1337-1	716	KEW3	34.59	-20.555	14.60	9.119	5.3
JBG3134	1001	KEW4	36.60	-16.822	15.23	8.362	28.6
JBG1724	1009	KEW5	39.11	-11.815	16.16	9.096	59.9
JBG1030	768	KEW6	34.49	-20.171	14.43	8.687	7.7
JBG975-1	N/A	KEW7	39.81	-14.861	16.42	8.532	40.9
JBG803-1	N/A	KEW8	41.94	-13.651	17.31	8.677	48.4
JBG726-1	951	KEW9	32.78	-18.852	13.61	8.944	15.9
JBG635-1	1010	KEW10	35.34	-12.829	14.66	8.629	53.6
JBG274-1	701	KEW11	30.12	-20.263	12.53	9.652	7.1
JBG98-1	1013	KEW12	31.57	-19.314	13.01	9.8	13.0

Table 4.4: A summary of the stable isotope results for the apatite samples, as well as the apatite-collagen δ13C difference for each individual

Individual	Tissue	Sample ID	δ18O	Wt% C	$\delta^{13}C$	% C4 diet (whole diet)	Ap-Col $\delta^{13}C$ Difference
JBG6134-1	Bone	KEW13	27.002	1.11%	-4.987	75.7	6.937
	Enamel	KEW32	25.834	0.86%	-4.049	-	-
JBG5499	Bone	KEW14	26.386	1.46%	-4.662	77.7	6.573
	Enamel	KEW31	27.081	0.86%	-4.963	-	-
JBG1337-1	Bone	KEW15	26.575	1.04%	-11.188	37.0	9.367
JBG3134	Bone	KEW16	27.148	0.88%	-8.182	55.7	8.64
	Enamel	KEW30	25.253	0.70%	-8.234	-	-
JBG1724	Bone	KEW17	27.023	0.51%	-5.502	72.5	6.313
	Enamel	KEW29	29.590	0.91%	-4.605	-	-
JBG1030	Bone	KEW18	26.287	1.52%	-9.757	45.9	10.414
JBG975-1	Bone	KEW19	26.758	1.51%	-6.393	66.9	8.468
	Enamel	KEW28	22.813	0.61%	-6.131	-	-
JBG803-1	Bone	KEW20	26.845	0.81%	-7.866	57.7	5.785
JBG726-1	Bone	KEW21	25.488	1.11%	-10.561	40.9	8.291
	Enamel	KEW27	24.942	0.95%	-4.965	-	-
JBG635-1	Bone	KEW22	26.346	0.96%	-6.289	67.6	6.54
	Enamel	KEW26	25.458	0.77%	-8.007	-	-
JBG274-1	Bone	KEW23	26.427	0.48%	-11.979	32.0	8.284
JBG98-1	Bone	KEW24	25.950	0.63%	-10.475	41.4	8.839
	Enamel	KEW25	24.582	0.93%	-2.456	-	-

Table 4.5: A list of Mississippian and Late Woodland populations used to compare to the dog bone isotope values. The age of each sample set is listed, as well as the means and standard deviations for the collagen δ13C and δ15N values, and the reference for each dataset.

Site	**Time Period (AD)**	**δ13C**		**δ15N**		**Reference**
		Mean	Stdev	Mean	Stdev	
Knight	400	-19.7	2.5	10.9	0.8	(Rose, 2008)
Joe Gay	600-700	-19.5	1.9	9.9	0.2	(Rose, 2008)
Schild	870	-19.7	1.2	9.6	0.7	(Rose, 2008)
Yokem	1000	-15.6	3.2	10.6	0.5	(Rose, 2008)
Mound 72 high status	1000-1050	-17.2	1.8	9.1	1.1	(Ambrose et al., 2003)
Mound 72 low status	1000-1050	-16.9	2.2	8.3	0.4	(Ambrose et al., 2003)
Moundville	1050-1250	-15.1	4	7	2	(Schoeninger and Schurr, 1998)
Schild Knoll A	1100-1310	-14.13	2.83	8.74	0.74	(Schober, 1998)
Ontario	1100-1400	-13.4	2.2	12.1	0.4	(Katzenberg, 1989)
Material Services Quarry	1162-1295	-12.5	0.94	9.7	0.41	(Emerson et al., 2005)
Gentleman Farm	1162-1295	-11.7	1.04	9.5	0.42	(Emerson et al., 2005)
East St Louis Quarry	1200-1275	-10.96	1.1	9.05	0.64	(Hedman et al., 2002)
Florence Street	1200-1275	-11.24	0.96	9.87	0.44	(Hedman et al., 2002)
Range Burial Group 1	1200-1275	-11.37	1.5	9.43	0.83	(Hedman et al., 2002)
Corbin Mounds	1200-1275	-12.08	1.58	9.12	0.35	(Hedman et al., 2002)
Hill Prairie Mound	1200-1275	-14.57	3.88	9.65	0.57	(Hedman et al., 2002)
Upland Sites	1250-1275	-12.9	2.7	9.3	0.5	(Bender et al., 1981)
American Bottom	1250-1275	-11.7	2	9.3	0.6	Bender 1981
Yokem	1290	-14.6	2.1	10.6	0.9	(Rose, 2008)
Angel Mounds	1350-1450	-9	1.1	8.3	0.7	(Schurr and Schoeninger, 1995)
Wickliffe	1350-1450	-9.5	0.8	8.2	0.5	(Schurr and Schoeninger, 1995)
Moundville	1400-1500	-10.3	1.1	8.3	1	(Schoeninger and Schurr, 1998)
Ontario	1600-1650	-12.5	1	12.1	0.9	(Katzenberg, 1989)

Table 4.6: A list of species identified from Late Woodland and Mississippian sites, organized by type of organism. These species have all been identified from archaeological sites based on skeletal remains, seeds, pollen, nut shells, or other remnants. The scientific and common names are listed. *: Over 80 species of freshwater mussels can be found in Illinois, and they cannot be identified based on shell alone (the primary source of mussel archaeological remains), so specific species information for these mussels is unavailable.

Plants			
Fruits		*Hordeum pusillum*	Little Barley
Asimina triloba	Pawpaw	*Iva annua*	Sumpweed
Morus rubra	Mulberry	*Nicotiana spp.*	Tobacco
Prunus spp.	Plum/Cherry	*Phalaris caroliniana*	Maygrass
Rhus spp.	Sumac	*Polygonium erectum*	Erect Knotweed
Vitis riparia	Wild grape	*Zea mays*	Maize
Nuts		**Miscellaneous Wild Plants**	
Carya spp.	Hickory	*Amaranthus spp.*	Amaranth
Corylus americana	Hazelnut	*Desmodium spp.*	Tick Clover
Juglans nigra	Walnut	*Eleusine indica*	Goosegrass
Quercus spp.	Acorn	*Ipomoea spp.*	Wild Morning Glory
Domesticated Crop Plants		*Panicum spp.*	Panic Grass
Chenopodium berlandieri	Goosefoot	*Phaseolus vulgaris*	Wild Bean
Cucurbita pepo	Squash	*Portulaca spp.*	Purslane
Helianthus annuus	Sunflower	*Solanum spp.*	Nightshade
Animals			
Mammals		*Mergus spp.*	Merganser
Castor canadensis	Beaver	*Oxyura jamaicensis*	Ruddy Duck
Cervus canadensis	Elk	*Tringa spp.*	Yellowlegs
Geomys bursarius	Pocket gopher	*Tympanuchus cupido*	Prairie Chicken
Odocoileus virginianus	White-Tailed Deer	*T. phasianellus*	Sharp-Tailed Grouse
Ondatra zibethicus	Muskrat	**Fish**	
Procyon lotor	Raccoon	*Dorosoma cepedianum*	Gizzard Shad
Sciurus spp.	Squirrel	*Ictalurus punctatus*	Channel Catfish
Sylvilagus floridanus	Rabbit	*Ameiurus spp.*	Bullhead
Birds		*Family Ictaluridae*	Catfish
Aix sponsa	Wood Duck	*Ictiobus cyprinellus*	Bigmouth Buffalo
Anas acuta	Pintail	*Percopsis omiscomaycus*	Perch
Anas americana	Widgeon	*Sander vitreus*	Walleye
Anas platyrhynchos	Mallard	*Hybognathus spp.*	Minnow
Anas spp.	Teal	Family Catostomidae	Suckers
Antigone canadensis	Sandhill Crane	*Lepomis spp.*	Sunfish
Aythya americana	Redhead	*Micropterus spp.*	Bass
Aythya collaris	Ring-Necked Duck	*Esox spp.*	Pike
Branta canadensis	Canada Goose	*Aplodinotus grunniens*	Freshwater Drum
Bucephala spp.	Goldeneye	*Atractosteus spatula*	Gar
Chen caerulescens	Snow Goose	*Amia calva*	Bowfin

Table 4.6 (cont'd)

Family Podicipedidae	Grebe	*Acipenser fulvescens*	Sturgeon
Family Rallidae	Rail	**Miscellaneous**	
Fulica americana	American Coot	Class Bivalvia*	Freshwater Mussels
Meleagris gallopavo	Wild Turkey		

Table 4.7: A summary of taxa identified from shotgun sequencing of the coprolites that are native to Southern Illinois

Genus	Common Name	Number of individuals	Number of reads	Reads in control
Zea	Maize	3	28	0
Toxocara	Dog Roundworm	2	27	0
Canis	Dog	3	21	3
Solanum	Nightshade	6	14	0
Nicotiana	Tobacco	5	10	0
Ardea	Great Heron	4	7	1
Cucurbita	Squash	3	6	0
Lepomis	Freshwater Sunfish	1	5	0
Dorosoma	Gizzard Shad	2	2	0
Couesius	Lake Chub	1	1	0

Table 4.8: A list of genera identified from a BLAST search of results that matched the custom chloroplast and mitochondrial databases searched against the nt database from NCBI. The genus is listed, along with the common name, as well as whether the genus has historically been found in Illinois or at Late Woodland or Mississippian sites in the American Bottom.

Genus	Common Name	Found in Illinois?
Amaranthus	Amaranth	Local plant, identified from Mississippian archaeological site
Amborella	Amborella (New Caledonian shrub)	Non-local
Annona	Pawpaw	Local plant, identified from Mississippian archaeological site
Canis	Dog/Wolf	Species analyzed
Capsicum	Peppers	Possible crop plant
Corylus	Hazelnut	Local plant, identified from Mississippian archaeological site
Cucurbita	Squash	Local Late Woodland/Mississippian crop plant
Cytinus	Parasitic plant	Non-local
Digitaria	Crabgrass	Local plant
Diospyros	Persimmon and Ebony	Local Plant
Dorosoma	Gizzard Shad	Local species
Erythranthe	Monkey flower	Local plant
Helianthus	Sunflower	Local Late Woodland/Mississippian crop plant
Ipomoea	Morning glory	Local plant, identified from Mississippian archaeological site
Phoenix	Palm plant	Non-local
Pinus	Pine	Local plant
Sesamum	Sesame	Non-local
Solanum	Nightshade	Local plant, identified from Mississippian archaeological site
Toxocara	Parasitic Nematode	Possible parasite
Vitis	Grape	Local plant, identified from Mississippian archaeological site

References

Alaeddini, R., 2012. Forensic implications of PCR inhibition--A review. Forensic Science International Genetics. 6, 297–305.

Allentoft, M.E., Collins, M., Harker, D., Haile, J., Oskam, C.L., Hale, M.L., Campos, P.F., Samaniego, J.A., Gilbert, M.T.P., Willerslev, E., Zhang, G., Scofield, R.P., Holdaway, R.N., Bunce, M., 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proceedings of the Royal Society B: Biological Sciences. 279, 4724–33.

Allitt, S., Stewart, R.M., Messner, T., 2009. The Utility of Dog Bone (*Canis Familiaris*) in Stable Isotope Studies for Investigating the Presence of Prehistoric Maize (*Zea mays ssp. mays*): A Preliminary Study. North American Archaeologist. 29, 343–367.

Alonso, A., Martín, P., Albarrán, C., García, P., García, O., de Simón, L.F., García-Hirschfeld, J., Sancho, M., de La Rúa, C., Fernández-Piqueras, J., 2004. Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. Forensic Science international. 139, 141–9.

Ambrose, S., 1987. Chemical and isotopic techniques of diet reconstruction in eastern North America. In: Emergent Horticultural Economies of the Eastern Woodlands. pp. 87-107.

Ambrose, S., 1990. Preparation and characterization of bone and tooth collagen for isotopic analysis. Journal of Archaeological Science. 17, 431–451.

Ambrose, S., 1993. Isotopic analysis of paleodiets: methodological and interpretive considerations. In: Investigations of Ancient Human Tissue: Chemical Analyses in Anthropology. pp. 59-130.

Ambrose, S., Norr, L., 1993. Experimental evidence for the relationship of the carbon isotope ratios of whole diet and dietary protein to those of bone collagen and carbonate. In: Prehistoric Human Bone: Archaeology at the Molecular Level. pp. 1-37.

Ambrose, S.H., Buikstra, J., Krueger, H.W., 2003. Status and gender differences in diet at Mound 72, Cahokia, revealed by isotopic analysis of bone. Journal of Anthropological Archaeology. 22, 217–226.

Ambrose, S.H., Butler, B.M., Hanson, D.B., Hunter-Anderson, R.L., Krueger, H.W., 1997. Stable isotopic analysis of human diet in the Marianas Archipelago, western Pacific. American Journal of Physical Anthropology. 104, 343–361.

Ambrose, S.H., Katzenberg, M.A., 2000. Biogeochemical approaches to paleodietary analysis.

Balasse, M., Ambrose, S.H., Smith, A.B., Price, T.D., 2002. The Seasonal Mobility Model for Prehistoric Herders in the South-western Cape of South Africa Assessed by Isotopic Analysis of Sheep Tooth Enamel. Journal of Archaeological Science. 29, 917–932.

Balasse, M., Smith, A.B., Ambrose, S.H., Leigh, S.R., 2003. Determining Sheep Birth Seasonality by Analysis of Tooth Enamel Oxygen Isotope Ratios: The Late Stone Age Site of Kasteelberg (South Africa). Journal of Archaeological Science. 30, 205–215.

Bender, M.M., Baerreis, D.A., Steventon, R.L., 1981. Further Light on Carbon Isotopes and Hopewell Agriculture. American Antiquity. 46, 346–353.

Bocherens, H., Koch, P.L., Mariotti, A., Geraads, D., Jaeger, J., 1996. Isotopic biogeochemistry (13C, 18O) of mammalian enamel from African pleistocene homonid sites. Paleios. 11, 306–318.

Bon, C., Berthonaud, V., Maksud, F., Labadie, K., Poulain, J., Artiguenave, F., Wincker, P., Aury, J.-M., Elalouf, J.-M., 2012. Coprolites as a source of information on the genome and diet of the cave hyena. Proceedings of the Royal Society B: Biological Sciences. 279, 2825–2830.

Borgic, Q.L., Galloy, J.M., 2004. Domesticated Dog Remains from the Janey B. Goode Site. In: Joint Meetings of the Southeastern Archaeological Conference and the Midwestern Archaeological Conference, St. Louis, Missouri, October (Vol. 2).

Bowen, G.J., Ehleringer, J.R., Chesson, L.A., Thompson, A.H., Podlesak, D.W., Cerling, T.E., 2009. Dietary and physiological controls on the hydrogen and oxygen isotope ratios of hair from mid-20th century indigenous populations. American Journal of Physical Anthropology. 139, 494–504.

Buikstra, J.E., Milner, G.R., 1989. Isotopic and Archaeological Interpretations of Diet in the Central Mississippi Valley. Journal of Archaeological Science. 319–329.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST plus: architecture and applications. BMC Bioinformatics. 10, 1.

Cantwell, A., 1980. Middle Woodland dog ceremonialism in Illinois. The Wisconsin Archeologist. 61, 480–496.

Cook, R.A., Price, T.D., 2015. Maize, mounds, and the movement of people: Isotope analysis of a Mississippian/Fort Ancient region. Journal of Archaeological Science. 61, 112–128.

Cook, R.A., Schurr, M.R., 2009. Eating between the lines: Mississippian migration and stable carbon isotope variation in fort ancient populations. American Anthropologist. 111, 344–359.

Crawford, G.W., Smith, D.G., Bowyer, V.E., 1997. Dating the Entry of Corn (Zea Mays) into the Lower Great Lakes Region. American Antiquity. 62, 112–119.

Dabney, J., Meyer, M., Paabo, S., 2013. Ancient DNA Damage. Cold Spring Harbor Perspectives in Biology. 5, a012567–a012567.

Despommier, D., 2003. Toxocariasis : Clinical Aspects , Epidemiology , Medical Ecology , and Molecular Aspects. Clinical Microbiology Reviews. 16, 265–272.

Emerson, T.E., 1997. Cahokia and the Archaeology of Power. The University of Alabama Press, Tuscaloosa, Alabama.

Emerson, T.E., Hedman, K.M., Simon, M.L., 2005. Marginal horticulturalists or maize agriculturalists? Archaeobotanical, paleopathological, and isotopic evidence relating to Langford tradition maize consumption. Midcontinental Journal of Archaeology. 30, 67–118.

Fizet, M., Mariotti, A., Bocherens, H., Lange-Badré, B., Vandermeersch, B., Borel, J.P., Bellon, G., 1995. Effect of diet, physiology and climate on carbon and nitrogen stable isotopes of collagen in a late pleistocene anthropic palaeoecosystem: Marillac, Charente, France. Journal of Archaeological Science. 22, 67–79.

Fortier, A.C., 2015. Preliminary Analyses of Prehistoric Dog Feces (Coprolites) from the Janey B. Goode Site, St. Clair County, Illinois. Illinois Antiquity. 50, 16–18.

Fortier, A.C., Jackson, D.K., 2000. The formation of a Late Woodland heartland in the American Bottom, Illinois cal AD 650–900. In: Late Woodland Societies: Tradition and Tranformation across the Midcontinent. University of Nebraska Press, pp. 123–147.

Galloy, J.M., 2010. PART V : WOODLAND PERIOD The Janey B . Goode Site (11S1232): Highlights of Investigations at a Massive Late Prehistoric Site in the American Bottom. Illinois Archaeology. 22, 529–552.

Geyh, M.A., Grosjean, M., Núñez, L., Schotterer, U., 1999. Radiocarbon Reservoir Effect and the Timing of the Late-Glacial/Early Holocene Humid Phase in the Atacama Desert (Northern Chile). Quaternary Research. 52, 143–153.

Gilbert, M.T.P., Hansen, A.J., Willerslev, E., Rudbeck, L., Barnes, I., Lynnerup, N., Cooper, A., 2003. Characterization of genetic miscoding lesions caused by postmortem damage. American Journal of Human Genetics. 72, 48–61.

Guiry, E.J., 2012. Dogs as Analogs in Stable Isotope-Based Human Paleodietary Reconstructions: A Review and Considerations for Future Use. Journal of Archaeological Method and Theory. 19, 351–376.

Hart, J.P., 1999. Maize Agriculture Evolution in the Eastern Woodlands of North America : A Darwinian Perspective. Journal of Archaelogical Method and Theory. 6, 137–180.

Hedges, R.E.M., Reynard, L.M., 2007. Nitrogen isotopes and the trophic level of humans in archaeology. Journal of Archaeological Science. 34, 1240–1251.

Hedman, K., Hargrave, E. A., Ambrose, S.H., 2002. Late Mississippian Diet in the American Bottom: Stable Isotope Analyses of Bone Collagen and Apatite. Midcontinental Journal of Archaeology. 27, 237–271.

Hofreiter, M., Betancourt, J.L., Sbriller, A.P., Markgraf, V., McDonald, H.G., 2003. Phylogeny, diet, and habitat of an extinct ground sloth from Cuchillo Cura, Neuquen Province, southwest Argentina. Quaternary Research. 59, 364–378.

Hofreiter, M., Poinar, H.N., Spaulding, W.G., Bauer, K., Martin, P.S., Possnert, G., Pääbo, S., 2000. A molecular analysis of ground sloth diet through the last glaciation. Molecular Ecology. 9, 1975–1984.

Hogue, S., 2003. Corn dogs and hush puppies: diet and domestication at two protohistoric farmsteads in Oktibbeha County, Mississippi. Southeastern Archaeology. 22, 185–195.

Huson, D.H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., Tappu, R., 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Computational Biology. 12, 1–12.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics. 29, 1682–4.

Katzenberg, M., 1989. Stable isotope analysis of archaeological faunal remains from southern Ontario. Journal of Archaeological Science. 16, 319–329.

Kelly, J.E., 2007. The context for the early evidence of the southeastern ceremonial complex at Cahokia. In: Southeastern Ceremonial Complex: Chronology Content Contest.

Kelly, L.S., 1997. Patterns of Faunal Exploitation at Cahokia. In: Pauketat, T.R., Emerson, T.E. (Eds.), Cahokia: Domination and Ideology in the Mississippian World. University of Nebraska Press, pp. 69–88.

Kemp, B.M., Monroe, C., Judd, K.G., Reams, E., Grier, C., 2014. Evaluation of methods that subdue the effects of polymerase chain reaction inhibitors in the study of ancient and degraded DNA. Journal of Archaeological Science. 42, 373–380.

Kidder, T.R., 1992. Timing and Consequences of the Introduction of Maize Agriculture in the Lower Mississippi Valley. North American Archaeologist. 13, 15–41.

Lapham, H.A., 2010. A Baumer Phase Dog Burial from the Kincaid Site in Southern Illinois. Illinois Archaeology. 22, 437–463.

Leonard, J. A., Shanks, O., Hofreiter, M., Kreuz, E., Hodges, L., Ream, W., Wayne, R.K., Fleischer, R.C., 2007. Animal DNA in PCR reagents plagues ancient DNA research. Journal of Archaeological Science. 34, 1361–1366.

Lopinot, N.H., 1997. Cahokian Food Production Reconsidered. In: Pauketat, T.R., Emerson, T.E. (Eds.), Cahokia: Domination and Ideology in the Mississippian World. University of Nebraska Press.

Noe-Nygaard, N., 1988. δ 13 C-values of dog bones reveal the nature of changes in man's food resources at the mesolithic-neolithic transition, Denmark. Chemical Geology: Isotope Geoscience section. 13, 87–96.

Park, R., Epstein, S., 1960. Carbon isotope fractionation during photosynthesis. Geochimica et Cosmochimica Acta. 21, 110–126.

Parmalee, P.W., Paloumpis, A.A., Wilson, N., 1972. Animals Utilized by Woodland Peoples Occupying the Apple Creek Site, Illinois, Research Papers 5. Springfield, IL.

Pauketat, T.R., 2004. Ancient Cahokia and the Mississippians. Cambridge University Press.

Pedersen, M.W., Ruter, A., Schweger, C., Friebe, H., Staff, R.A., Kjeldsen, K.K., Mendoza, M.L.Z., Beaudoin, A.B., Zutter, C., Larsen, N.K., Potter, B.A., Nielsen, R., Rainville, R.A., Orlando, L., Meltzer, D.J., Kjær, K.H., Willerslev, E., 2016. Postglacial viability and colonization in North America's ice-free

corridor. Nature. 537, 45-49.

Philippsen, B., 2013. The freshwater reservoir effect in radiocarbon dating. Heritage Science. 1, 24.

Richardson, D.J., Guillén, S., Beckett, R., Kyle, W., Conlogue, G., Harper-Beckett, K., 2012. Archaeohelminthology of the Chiribaya Shepherd, Canis familiaris (700–1476 AD) from Southern Peru. Comparative Parasitology. 79, 133–137.

Riley, T.J., Walz, G.R., Bareis, C.J., Fortier, A.C., Parker, K.E., 1994. Accelerator Mass Spectrometry (AMS) Dates Confirm Early Zea Mays in the Mississippi River Valley. American Antiquity. 59, 490–498.

Robertson, I.D., Thompson, R.C., 2002. Enteric parasitic zoonoses of domesticated dogs and cats. Microbes and Infection. 4, 867–873.

Rose, F., 2008. Intra-Community Variation in Diet during the Adoption of a New Staple Crop in the Eastern Woodlands. American Antiquity. 73, 413–439.

Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biology. 12, 87.

Schober, T., 1998. Reinvestigation of Maize Introduction in West-Central Illinois. University of Illinois at Urbana Champaign. Master's Thesis.

Schoeninger, M.J., 2009. Stable isotope evidence for the adoption of maize agriculture. Current Anthropology. 50, 633–640.

Schoeninger, M.J., DeNiro, M.J., 1984. Nitrogen and carbon isotopic composition of bone collagen from marine and terrestrial animals. Geochimica et Cosmochimica Acta. 48, 625–639.

Schoeninger, M.J., Schurr, M.R., 1998. Human Subsistence at Moundville: the Stable-Isotope Data. In: Knight, V.J.J., Steponaitis, V.P. (Eds.), Archaeology of the Moundville Chiefdom. Smithsonian Institution Press, Washington DC, pp. 120–132.

Schroeder, S., 1999. Maize productivity in the eastern woodlands and Great Plains of North America. American Antiquity. 64, 499–516.

Schurr, M., Schoeninger, M.J., 1995. Associations between agricultural intensification and social complexity: An example from the Prehistoric Ohio Valley. Journal of Anthropological Archaeology. 14, 315–339.

Schwarcz, H.P., 1991. Some theoretical aspects of isotope paleodiet studies. Journal of Archaeological Science. 18, 261–275.

Schwartz, M., 2000. The Form and Meaning of Maya and Mississippian Dog Representations. In: Dogs Through Time: An Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. pp. 217–226.

Schwarz, C., Debruyne, R., Kuch, M., McNally, E., Schwarcz, H., Aubrey, A.D., Bada, J., Poinar, H., 2009.

New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. Nucleic Acids Research. 37, 3215–29.

Simon, M., 2010. Beyond Gathering : Early Plant Cultivation in Illinois. Illinois Antiquity. 45, 17–21.

Simon, M.L., 2017. Reevaluating the Evidence for Middle Woodland Maize From the Holding Site. American Antiquity. 82, 140–150.

Simon, M.L., Lopinot, K.E., 2006. PREHISTORIC PLANT USE IN THE AMERICAN BOTTOM : NEW THOUGHTS AND INTERPRETATIONS. Southeastern Archaeology. 25, 212–257.

Simpson, J.T., Wong, K., Jackman, S.D., Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., 2009. ABySS : A parallel assembler for short read sequence data ABySS : A parallel assembler for short read sequence data. 1117–1123.

Smith, B., 1989. Origins of agriculture in eastern North America. Science. 246, 1566–1571.

Smith, B.D., 2011. The Cultural Context of Plant Domestication in Eastern North America. Current Anthropology. 52, S471–S484.

Sponheimer, M., Lee-Thorp, J.A., 1999. Oxygen Isotopes in Enamel Carbonate and their Ecological Significance. Journal of Archaeological Science. 26, 723–728.

Tankersley, K.B., Koster, J.M., 2009. Sources of Stable Isotope Variation in Archaeological Dog Remains. North American Archaeologist. 30, 361–375.

Tykot, R.H., van der Merwe, N.J., Hammond, N., 1996. Stable isotope analysis of bone collagen, bone apatite, and tooth enamel in the reconstruction of human diet: a case study from Cuello, Belize. In: Archaeological chemistry: organic, inorganic, and biochemical analysis. pp. 355–365.

van der Merwe, N.J., Vogel, J.C., 1978. 13C content of human collagen as a measure of prehistoric diet in woodland North America. Nature. 276, 815-816.

VanDerwarker, A.M., Bardolph, D.N., Hoppa, K.M., Thakar, H.B., Martin, L.S., Jaqua, A.L., Biwer, M.E., Gill, K.M., 2016. New World Paleoethnobotany in the New Millennium (2000-2013). Journal of Archaeological Research. 24, 125–177.

Vanderwarker, A.M., Wilson, G.D., Bardolph, D.N., 2013. Maize Adoption and Intensification in the Central Illinois River Valley: An Analysis of Archaeobotanical Data from the Late Woodland to Early Mississippian Periods (A.D. 600-1200). Southeastern Archaeology. 32, 147–168.

Warinner, C., Rodrigues, J.F.M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R.Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H.U., Salazar-García, D.C., Eppler, E., Seiler, R., Hansen, L.H., Castruita, J.A.S., Barkow-Oesterreicher, S., Teoh, K.Y., Kelstrup, C.D., Olsen, J. V, Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis, C.M., Collins, M.J., Gilbert, M.T.P., Rühli, F., Cappellini, E., 2014. Pathogens and host immunity in the ancient human oral cavity. Nature Genetics. 46, 336–44.

Welker, F., Duijm, E., van der Gaag, K.J., van Geel, B., de Knijff, P., van Leeuwen, J., Mol, D., van der

Plicht, J., Raes, N., Reumer, J., Gravendeel, B., 2014. Analysis of coprolites from the extinct mountain goat Myotragus balearicus. Quaternary Research. 81, 106–116.

White, C., Longstaffe, F.J., Law, K.R., 2004. Exploring the effects of environment, physiology and diet on oxygen isotope ratios in ancient Nubian bones and teeth. Journal of Archaeological Science. 31, 233–250.

White, C.D., Pohl, M.E.D., Schwarcz, H.P., Longstaffe, F.J., 2001. Isotopic Evidence for Maya Patterns of Deer and Dog Use at Preclassic Colha. Journal of Archaeological Science. 28, 89–107.

Willerslev, E., Cooper, A., 2005. Ancient DNA. Proceedings of the Royal Society B: Biological Sciences. 272, 3–16.

Wood, J.R., Rawlence, N.J., Rogers, G.M., Austin, J.J., Worthy, T.H., Cooper, A., 2008. Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). Quaternary Science Reviews. 27, 2593–2602.

Wright, L.E., Schwarcz, H.P., 1999. Correspondence Between Stable Carbon, Oxygen and Nitrogen Isotopes in Human Tooth Enamel and Dentine: Infant Diets at Kaminaljuyú. Journal of Archaeological Science. 26, 1159–1170.

Yerkes, R.W., 2011. Bone Chemistry, Body Parts, and Growth Marks: Evaluating Ohio Hopewell and Cahokia Mississippian Seasonality, Subsistence, Ritual, and Feasting. American Archaeology. 70, 241–265.

CHAPTER 5: CONCLUSIONS

Summary of Findings

The domestic dog has experienced a unique and varied relationship with humans for thousands of years, and the dog genome reflects generations of breeding for specific traits and adaptation to new environments (Lindblad-Toh et al., 2005; Karlsson et al., 2007). The history of dogs has been useful for inferring the history of humans, both in terms of understanding population turnover and continuity (Brown et al., 2013; Greig et al., 2015), and in terms of examining how humans (and dogs) adapt to new lifestyles and environments (Axelsson et al., 2013; Li et al., 2014). When humans peopled the Americas, they brought dogs with them (Leonard et al., 2002), and many groups depended on their dogs and valued them highly (Schwartz, 1997). Because of this connection between humans and dogs in the Americas, we can learn more about human history through the study of their dogs.

In Chapter 2, through the analysis of the hypervariable region of mitochondrial DNA from several ancient dog populations, I identified multiple founding haplotypes, which were found in dogs across the Americas. One haplotype was particularly common, and the haplotype network shows a star-like pattern, suggesting that the dog population expanded rapidly following the arrival to the Americas. Some of the dogs in the Americas had haplotypes that only differed from wolf haplotypes by two to three nucleotides, suggesting that there may have been some admixture with wolves in the Arctic. I also found that dogs from South America clustered with one another, separate from dogs from North America, which is similar in phylogeographic terms to the Northern and Southern clades found in Native American populations (Rasmussen et al., 2014; Verdu et al., 2014).

Dog populations also vary in their levels of genetic diversity. This could reflect a difference in founding population size, bottlenecks due to human cultural changes or natural environmental events, or differences in breeding practices. Demographic modeling of the populations suggests that dogs may have been introduced to the Americas as recently as 10,000 ybp, which is consistent with our findings in the

archaeological record (Morey and Wiant, 1992; Walker et al., 2005). Humans arrived in multiple migration "waves" (Raghavan et al., 2014), and so it seemed possible that dogs were a later introduction to the Americas.

In chapter 3, analysis of complete mitochondrial genomes, however, challenges some previous assumptions about dogs in the Americas. For example, although previous studies have identified ancient haplotypes in modern dogs when comparing hypervariable region sequences (van Asch et al., 2013), comparison of mitogenome sequences indicated that ancient haplotypes are not found in modern dogs. The loss of Native American genetic diversity as a result of European colonization has been previously shown (O'Fallon and Fehren-Schmitz, 2011; Lindo et al., 2016; Llamas et al., 2016), and dogs seem to have undergone a similar bottleneck, although it began before European colonization, likely as a result of the changing role of dogs through time across the Americas. In the case of the dogs, the dog population shows a decline beginning 2000 ybp. One disadvantage of using mitogenome sequences is that the majority of published modern dog mitogenome sequences are from breed dogs, which are European in origin. There may be some indeterminate breed dogs that still harbor ancient American haplotypes, but they have not yet been sequenced.

Demographic modeling of the complete mitochondrial genome also challenges the hypothesis that dogs may have arrived to the Americas after the initial migration, which was suggested in Chapter 2. Instead, the modeling of the complete mitogenome suggests that dogs arrived to the Americas between 17,000 and 13,000 ybp, which is consistent with the timing of the initial peopling of the Americas (Kemp and Schurr, 2010; Meltzer, 2010; Llamas et al., 2016). Dogs in the Americas show divergence from Siberian dogs, which suggests a long period of isolation between the populations prior to entering the Americas. This is consistent with Native American populations, who show a similar divergence from their Siberian source population (Tamm et al., 2007).

Some of the findings from the mitogenome data supported what was known from the HVR

sequences. For example, dog populations showed varying levels of genetic diversity, and Northern and Southern mitochondrial DNA clades of dogs were identified. One clade included dogs from North America as well as the Yucatan and Argentina, while the other was more localized to the Southeast and Midwestern United States. Additionally, as previously mentioned, the genetic similarity between ancient American dogs and Eurasian wolves supported the hypothesis that dogs migrated with humans into the Americas, rather than being domesticated separately.

Most interestingly, using dog mitogenome sequences, we identified shared genetic lineages between dog populations that mirror those in human populations. For example, we showed relatedness between dogs along the Pacific coast, which is consistent with a human migration route down the Pacific coast (Eshleman et al., 2004; Wang et al., 2007). We also found that dogs from the Midwest and Southeast were closely related, which is similar to geographic patterns found across human populations from the same region, who share a number of cultural connections from the Woodland and Mississippian periods (Seeman, 1979; Wolfe Steadman, 2001; Pauketat and Alt, 2015). And with the large depth of sampling in the Midwest, we were able to demonstrate that the dog population there increased in genetic diversity over time. The increase between 1200 and 800 ybp coincides with the Late Woodland - Mississippian transition, suggesting that dogs were introduced from other regions of the Americas as people migrated to be closer to Cahokia, or as trade routes between Cahokia and other regions of the Americas expanded. The study of dogs can thus yield a wealth of information about human populations as well.

In Chapter 4, I aimed to assess the timing of maize consumption of humans in the American Bottom by using dogs as a dietary proxy. Around the time of the arrival of maize to Southern Illinois, there was also a large cultural change in the area, with the start of the Mississippian period (Kelly, 1990; Emerson, 1997; Pauketat and Emerson, 1997; Smith, 2007). Maize was long thought to be an important part of Mississippian diet, but its level of importance to Mississippian culture has been debated (Kelly, 1990; Lopinot, 1997; Vanderwarker et al., 2013). The study of dietary isotopes demonstrates that dogs

from the Late Woodland period had little to no maize in their diets, while dogs from the Mississippian period dating to 1010 AD and later were eating large amounts of maize. These changes in diet are similar to what's been identified in humans from other sites from the Midwestern United States (van der Merwe and Vogel, 1978; Hedman et al., 2002; Emerson et al., 2005; Yerkes, 2011). Maize was identified in three of the coprolites analyzed, demonstrating that dogs were eating maize directly. Radiocarbon dating these coprolites can further clarify when dogs, and by proxy, humans, began to eat maize.

In addition to maize, we also examined other aspects of the dogs' diet. The $\delta^{13}C$ of the dietary protein was lower than the $\delta^{13}C$ of the whole diet, which suggests that these dogs were not consuming maize only. A number of fish species were identified from the coprolites, including lake chub and gizzard shad, and some of these were identified from the same coprolites that had maize, demonstrating that fish was another likely source of protein. Other native species that were identified include herons, as well as crops like squash and tobacco. Dogs are often a reliable indicator of human diet (Guiry, 2012), and so this suggests that Mississippian humans consumed a varied diet of maize, fish, and other crops that were first cultivated during the Late Woodland period.

Future Directions

Although this research has yielded much detail of the history of dogs in the Americas, as well as how they interacted with humans, additional questions remain regarding the demographic history of dogs. These questions can be addressed by expanding the sampling of dogs in the Americas, by sampling from additional archaeological sites and by sequencing additional regions of the genome. By expanding the geographic sampling, we can get a fuller picture of the history of dogs in the Americas. Only a small number of geographic regions were sampled in these studies, including the Midwest, Southwest, Pacific Coast, and Southeastern United States, as well as the Yucatan. Most of South America, Canada, and Mexico have had limited sampling of ancient dogs, despite the presence of dog remains, especially in North America (Morey, 2006). Additionally, regions like the Great Plains and the Northeastern United

States have a wealth of dog burials that could be used to construct population histories similar to the coverage of the Midwest (Walker and Frison, 1982; Schwartz, 1997; Yohe and Pavesic, 2000; Morey, 2006). By sampling dogs from multiple archaeological sites and time periods from the same region, it will be possible to reconstruct the history of dogs in an area. Changes to dog demography in a region can reflect human cultural population changes – a population bottleneck could show an intensification of deliberate breeding, or an increase in diversity could indicate the introduction of new haplotypes, as a result of human migration or trade.

While mitochondrial DNA is useful as an indicator of ancestry, other regions of the genome can reveal much more about ancient dogs. The only region of the genome that has been sequenced in ancient American dogs is the mitogenome, and so genomic sequencing should be attempted in ancient American dogs. Sequencing the complete genome can reveal traits that were under selection, and also reveal admixture in a population, and mitochondrial DNA data has limited utility in examining these aspects of demographic history. Numerous genes that code for phenotypic traits have been identified in dogs (Cadieu et al., 2009; Rimbault and Ostrander, 2012; Hayward et al., 2016), and these traits can give an idea of what the dogs may have looked like, or what role they played in human lives. Signatures of selection can be indicators of human-mediated selection (such as deliberate breeding for phenotypic traits), or of responses to new environments (such as adaptation to starch digestion). Mitochondrial DNA reflects only a fraction of genetic ancestry, and so by sequencing complete genomes, it would be possible to test hypotheses of dog and wolf admixture in the Americas, or of population continuity between ancient dogs and modern dogs.

The study of dog diet demonstrated that dogs (and probably humans) were eating maize during the early Mississippian period, but additional analysis can further our understanding of their diet. Even after filtering the results using criteria that were designed to screen out contaminants, and further filtering the results by searching them against a larger database, four genera remain in the dataset that aren't

native to Illinois. Further study of these results is needed to determine if these taxa are additional contaminants, or just the closest relative to a native species in the database. By sequencing DNA from local species, it would be possible to better compare the coprolite dataset to what species were known to live in the area.

Additionally, shotgun sequencing of coprolites has yielded millions of sequencing reads, many of which have been discarded in the dietary analysis. Fecal samples are a rich source of microbiome data, and can yield information about an individual's health, diet, and other factors. Dog gut microbiomes have been sequenced (Middelbos et al., 2010; Swanson et al., 2011), and it would be interesting to compare the microbiomes of modern and ancient dogs. Ancient microbiomes of humans have been successfully sequenced in the past, and some have yielded gut microbiota that are similar to those found in modern human gut microbiomes (Warinner et al., 2015). The differences between ancient and modern dog microbiomes might be due to diet, but they also may be due to adaptation to different environments, and comparing ancient and modern dog microbiomes can reveal more about the lives of ancient dogs.

With the limited availability of ancient remains, especially human remains, the use of other organisms can help fill in gaps of knowledge about human migration and cultural history. Dogs are especially appropriate for filling in those gaps, because they have such a close relationship with humans and have traveled with them so extensively. They can serve as genetic proxies, to reconstruct migration routes and population interactions, as well as dietary proxies, to show what past peoples ate and how they lived. People have valued dogs for centuries as hunters, guards, and companions, and as we uncover ancient dog remains, they can fill a new role: an archive of human history.

References

Axelsson, E., Ratnakumar, A., Arendt, M.-L., Maqbool, K., Webster, M.T., Perloski, M., Liberg, O., Arnemo, J.M., Hedhammar, A., Lindblad-Toh, K., 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature. 495, 360-364.

Brown, S.K., Darwent, C.M., Sacks, B.N., 2013. Ancient DNA evidence for genetic continuity in arctic dogs. Journal of Archaeological Science. 40, 1279–1288.

Cadieu, E., Neff, M.W., Quignon, P., Walsh, K., Chase, K., Parker, H.G., Vonholdt, B.M., Rhue, A., Boyko, A., Byers, A., Wong, A., Mosher, D.S., Elkahloun, A.G., Spady, T.C., André, C., Lark, K.G., Cargill, M., Bustamante, C.D., Wayne, R.K., Ostrander, E. a, 2009. Coat variation in the domestic dog is governed by variants in three genes. Science. 326, 150–3.

Emerson, T.E., 1997. Cahokia and the Archaeology of Power. The University of Alabama Press, Tuscaloosa, Alabama.

Emerson, T.E., Hedman, K.M., Simon, M.L., 2005. Marginal horticulturalists or maize agriculturalists? Archaeobotanical, paleopathological, and isotopic evidence relating to Langford tradition maize consumption. Midcontinental Journal of Archaeology. 30, 67–118.

Eshleman, J.A., Malhi, R.S., Johnson, J.R., Kaestle, F.A., Lorenz, J., Smith, D.G., 2004. Mitochondrial DNA and prehistoric settlements: native migrations on the western edge of North America. Human Biology. 76, 55–75.

Greig, K., Boocock, J., Prost, S., Horsburgh, K.A., Jacomb, C., Walter, R., Matisoo-Smith, E., 2015. Complete Mitochondrial Genomes of New Zealand's First Dogs. PLoS ONE. 10, e0138536.

Guiry, E.J., 2012. Dogs as Analogs in Stable Isotope-Based Human Paleodietary Reconstructions: A Review and Considerations for Future Use. Journal of Archaeological Method and Theory. 19, 351–376.

Hayward, J., Castelhano, M., Oliveira, K., Corey, E., Balkman, C., Baxter, T., Casal, M., Center, S., Fang, M., Garrison, S., Kalla, S., Korniliev, P., Kotlikoff, M., Moise, N., Shannon, L., Simpson, K., Sutter, N., Todhunter, R., Boyko, A., 2016. Complex disease and phenotype mapping in the domestic dog. Nature Communications. 7, 10460.

Hedman, K., Hargrave, E. A., Ambrose, S.H., 2002. Late Mississippian Diet in the American Bottom: Stable Isotope Analyses of Bone Collagen and Apatite. Midcontinental Journal of Archaeology. 27, 237–271.

Karlsson, E.K., Baranowska, I., Wade, C.M., Salmon Hillbertz, N.H.C., Zody, M.C., Anderson, N., Biagi, T.M., Patterson, N., Pielberg, G.R., Kulbokas, E.J., Comstock, K.E., Keller, E.T., Mesirov, J.P., von Euler, H., Kämpe, O., Hedhammar, A., Lander, E.S., Andersson, G., Andersson, L., Lindblad-Toh, K., 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. Nature Genetics. 39, 1321–1328.

Kelly, J.E., 1990. The Emergence of Mississippian Culture in the American Bottom Region. In: The Mississippian Emergence. pp. 113–152.

Kemp, B.M., Schurr, T.G., 2010. Ancient and modern genetic variation in the Americas. In: Auerbach, B. (Ed.), Human Variation in the Americas: The Integration of Archeology and Biological Anthropology. Carbondale, IL, pp. 12–50.

Leonard, J. a, Wayne, R.K., Wheeler, J., Valadez, R., Guillén, S., Vilà, C., 2002. Ancient DNA evidence for Old World origin of New World dogs. Science. 298, 1613–6.

Li, Y., Wu, D.-D., Boyko, A.R., Wang, G.-D., Wu, S.-F., Irwin, D.M., Zhang, Y.-P., 2014. Population Variation Revealed High-Altitude Adaptation of Tibetan Mastiffs. Molecular biology and evolution. 31, 1200–1205.

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E. a, Ponting, C.P., Galibert, F., Smith, D.R., DeJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.-W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.-P., Parker, H.G., Pollinger, J.P., Searle, S.M.J., Sutter, N.B., Thomas, R., Webber, C., Baldwin, J., Abebe, A., Abouelleil, A., Aftuck, L., Ait-Zahra, M., Aldredge, T., Allen, N., An, P., Anderson, S., Antoine, C., Arachchi, H., Aslam, A., Ayotte, L., Bachantsang, P., Barry, A., Bayul, T., Benamara, M., Berlin, A., Bessette, D., Blitshteyn, B., Bloom, T., Blye, J., Boguslavskiy, L., Bonnet, C., Boukhgalter, B., Brown, A., Cahill, P., Calixte, N., Camarata, J., Cheshatsang, Y., Chu, J., Citroen, M., Collymore, A., Cooke, P., Dawoe, T., Daza, R., Decktor, K., DeGray, S., Dhargay, N., Dooley, K., Dooley, K., Dorje, P., Dorjee, K., Dorris, L., Duffey, N., Dupes, A., Egbiremolen, O., Elong, R., Falk, J., Farina, A., Faro, S., Ferguson, D., Ferreira, P., Fisher, S., FitzGerald, M., Foley, K., Foley, C., Franke, A., Friedrich, D., Gage, D., Garber, M., Gearin, G., Giannoukos, G., Goode, T., Goyette, A., Graham, J., Grandbois, E., Gyaltsen, K., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Healy, C., Hegarty, R., Honan, T., Horn, A., Houde, N., Hughes, L., Hunnicutt, L., Husby, M., Jester, B., Jones, C., Kamat, A., Kanga, B., Kells, C., Khazanovich, D., Kieu, A.C., Kisner, P., Kumar, M., Lance, K., Landers, T., Lara, M., Lee, W., Leger, J.-P., Lennon, N., Leuper, L., LeVine, S., Liu, J., Liu, X., Lokyitsang, Y., Lokyitsang, T., Lui, A., Macdonald, J., Major, J., Marabella, R., Maru, K., Matthews, C., McDonough, S., Mehta, T., Meldrim, J., Melnikov, A., Meneus, L., Mihalev, A., Mihova, T., Miller, K., Mittelman, R., Mlenga, V., Mulrain, L., Munson, G., Navidi, A., Naylor, J., Nguyen, T., Nguyen, N., Nguyen, C., Nguyen, T., Nicol, R., Norbu, N., Norbu, C., Novod, N., Nyima, T., Olandt, P., O'Neill, B., O'Neill, K., Osman, S., Oyono, L., Patti, C., Perrin, D., Phunkhang, P., Pierre, F., Priest, M., Rachupka, A., Raghuraman, S., Rameau, R., Ray, V., Raymond, C., Rege, F., Rise, C., Rogers, J., Rogov, P., Sahalie, J., Settipalli, S., Sharpe, T., Shea, T., Sheehan, M., Sherpa, N., Shi, J., Shih, D., Sloan, J., Smith, C., Sparrow, T., Stalker, J., Stange-Thomann, N., Stavropoulos, S., Stone, C., Stone, S., Sykes, S., Tchuinga, P., Tenzing, P., Tesfaye, S., Thoulutsang, D., Thoulutsang, Y., Topham, K., Topping, I., Tsamla, T., Vassiliev, H., Venkataraman, V., Vo, A., Wangchuk, T., Wangdi, T., Weiand, M., Wilkinson, J., Wilson, A., Yadav, S., Yang, S., Yang, X., Young, G., Yu, Q., Zainoun, J., Zembek, L., Zimmer, A., Lander, E.S., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature. 438, 803–19.

Lindo, J., Huerta-Sanchez, E., Nakagome, S., Rasmussen, M., Petzelt, B., Mitchell, J., Cybulski, J.S., Willerslev, E., DeGiorgio, M., Malhi, R.S., 2016. A time transect of exomes from a Native American population before and after European contact. Nature Communications. 7, 13175.

Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S.M., Rohlach, A., Barreto-Romero, M.I., Flores-Espinoza, I., Tomasto-Cagigao, E., Watson-Jimenez, L., Makowski, K., Santiago-Leboreiro-Reyna, I., Mansilla-Lory, J.,

Ballivian-Torrez, J.A., Rivera, M.A., Burger, R.L., Constanza-Ceruti, M., Reinhard, J., Wells, R.S., Politis, G., Santoro, C.M., Standen, V.G., Smith, C., Reich, D., Ho, S.Y.W., Cooper, A., Haak, W., 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Science Advances. 2, e1501385.

Lopinot, N.H., 1997. Cahokian Food Production Reconsidered. In: Pauketat, T.R., Emerson, T.E. (Eds.), Cahokia: Domination and Ideology in the Mississippian World. University of Nebraska Press.

Meltzer, D., 2010. First Peoples in a New World: Colonizing Ice Age America. University of California Press.

Middelbos, I.S., Vester Boler, B.M., Qu, A., White, B. A., Swanson, K.S., Fahey, G.C., 2010. Phylogenetic characterization of fecal microbial communities of dogs fed diets with or without supplemental dietary fiber using 454 pyrosequencing. PloS ONE. 5, e9768.

Morey, D.F., 2006. Burying key evidence : the social bond between dogs and people. Journal of Archaeological Science. 33, 158–175.

Morey, D.F., Wiant, M.D., 1992. Early Holocene Domestic Dog Burials from the North American Midwest. Current Anthropology. 33, 224–230.

O'Fallon, B.D., Fehren-Schmitz, L., 2011. Native Americans experienced a strong population bottleneck coincident with European contact. Proceedings of the National Academy of Sciences of the United States of America. 108, 20444–8.

Pauketat, T.R., Alt, S.M., 2015. Medieval Mississippians: The Cahokian World.

Pauketat, T.R., Emerson, T.E., 1997. Cahokia: Domination and Ideology in the Mississippian World. University of Nebraska Press.

Raghavan, M., DeGiorgio, M., Albrechtsen, A., Moltke, I., Skoglund, P., Korneliussen, T.S., Gronnow, B., Appelt, M., Gullov, H.C., Friesen, T.M., Fitzhugh, W., Malmstrom, H., Rasmussen, S., Olsen, J., Melchior, L., Fuller, B.T., Fahrni, S.M., Stafford, T., Grimes, V., Renouf, M.A.P., Cybulski, J., Lynnerup, N., Lahr, M.M., Britton, K., Knecht, R., Arneborg, J., Metspalu, M., Cornejo, O.E., Malaspinas, A.-S., Wang, Y., Rasmussen, M., Raghavan, V., Hansen, T.V.O., Khusnutdinova, E., Pierre, T., Dneprovsky, K., Andreasen, C., Lange, H., Hayes, M.G., Coltrain, J., Spitsyn, V.A., Gotherstrom, A., Orlando, L., Kivisild, T., Villems, R., Crawford, M.H., Nielsen, F.C., Dissing, J., Heinemeier, J., Meldgaard, M., Bustamante, C., O'Rourke, D.H., Jakobsson, M., Gilbert, M.T.P., Nielsen, R., Willerslev, E., 2014. The genetic prehistory of the New World Arctic. Science. 345, 1255832.

Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir, V., Yadav, R., Malaspinas, A.-S., White, S.S., Allentoft, M.E., Cornejo, O.E., Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M., Korneliussen, T.S., Meltzer, D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S., Brunak, S., Sicheritz-Ponten, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica, A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., Willerslev, E., 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature. 506, 225–9.

Rimbault, M., Ostrander, E.A., 2012. So many doggone traits : mapping genetics of multiple phenotypes in the domestic dog. Human Molecular Genetics. 21, 52–57.

Schwartz, M., 1997. A History of Dogs in the Early Americas. Yale University Press, New Haven, CT.

Seeman, M.F., 1979. Hopewell Interaction Sphere: the Evidence for Interregional Trade and Structural Complexity.

Smith, B.D., 2007. The Mississippian Emergence. University of Alabama Press.

Swanson, K.S., Dowd, S.E., Suchodolski, J.S., Middelbos, I.S., Vester, B.M., Barry, K. A., Nelson, K.E., Torralba, M., Henrissat, B., Coutinho, P.M., Cann, I.K.O., White, B. A., Fahey, G.C., 2011. Phylogenetic and gene-centric metagenomics of the canine intestinal microbiome reveals similarities with humans and mice. The ISME journal. 5, 639–49.

Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D.G., Mulligan, C.J., Bravi, C.M., Rickards, O., Martinez-Labarga, C., Khusnutdinova, E.K., Fedorova, S. a, Golubenko, M. V, Stepanov, V. A., Gubina, M. A., Zhadanov, S.I., Ossipova, L.P., Damba, L., Voevoda, M.I., Dipierri, J.E., Villems, R., Malhi, R.S., 2007. Beringian standstill and spread of Native American founders. PloS ONE. 2, e829.

van Asch, B., Zhang, A.-B., Oskarsson, M.C.R., Klütsch, C.F.C., Amorim, A., Savolainen, P., 2013. Pre-Columbian origins of Native American dog breeds, with only limited replacement by European dogs, confirmed by mtDNA analysis. Proceedings of the Royal Society B: Biological Sciences. 280, 20131142.

van der Merwe, N.J., Vogel, J.C., 1978. 13C content of human collagen as a measure of prehistoric diet in woodland North America. Nature. 276, 815-816.

Vanderwarker, A.M., Wilson, G.D., Bardolph, D.N., 2013. Maize Adoption and Intensification in the Central Illinois River Valley: An Analysis of Archaeobotanical Data from the Late Woodland to Early Mississippian Periods (A.D. 600-1200). Southeastern Archaeology. 32, 147–168.

Verdu, P., Pemberton, T.J., Laurent, R., Kemp, B.M., Gonzalez-Oliver, A., Gorodezky, C., Hughes, C.E., Shattuck, M.R., Petzelt, B., Mitchell, J., Harry, H., William, T., Worl, R., Cybulski, J.S., Rosenberg, N.A., Malhi, R.S., 2014. Patterns of admixture and population structure in native populations of Northwest North America. PLoS Genetics. 10, e1004530.

Walker, D.N., Frison, G.C., 1982. Studies on Amerindian Dogs , 3 : Prehistoric Wolf / Dog Hybrids from the Northwestern Plains. Journal of Archaeological Science. 9, 125–172.

Walker, R., Morey, D., Relethford, J., 2005. Early and Mid-Holocene Dogs in Southeastern North America: Examples from Dust Cave. Southeastern Archaeology. 83-92.

Wang, S., Lewis, C.M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M. V, Molina, J. a, Gallo, C., Mazzotti, G., Poletti, G., Hill, K., Hurtado, A.M., Labuda, D., Klitz, W., Barrantes, R., Bortolini, M.C., Salzano, F.M., Petzl-Erler, M.L., Tsuneto, L.T., Llop, E., Rothhammer, F., Excoffier, L., Feldman, M.W., Rosenberg, N. a, Ruiz-Linares, A., 2007. Genetic variation and population structure in native Americans. PLoS Genetics. 3, e185.

Warinner, C., Speller, C., Collins, M.J., Lewis, C.M., 2015. Ancient human microbiomes. Journal of Human Evolution. 79, 125–136.

Wolfe Steadman, D., 2001. Mississippians in motion? A population genetic analysis of interregional gene flow in West-Central Illinois. American Journal of Physical Anthropology. 114, 61–73.

Yerkes, R.W., 2011. Bone Chemistry, Body Parts, and Growth Marks: Evaluating Ohio Hopewell and Cahokia Mississippian Seasonality, Subsistence, Ritual, and Feasting. American Archaeology. 70, 241–265.

Yohe, R., Pavesic, M., 2000. Early archaic domestic dogs from Western Idaho, USA. In: Dogs Through Time: An Archaeological Perspective. Proceedings of the 1st ICAZ Symposium on the History of the Domestic Dog. August 23-29, 1998. pp. 93–104.